Genotypic and phenotypic consequences of domestication in dogs

Authors: Sweetalana[1], Shirin Nataneli[2], Shengmiao Huang[2], Jazlyn A Mooney[2,*] and Zachary A Szpiech[1,3,*]

[1]Department of Biology, Pennsylvania State University, University Park, PA, USA

[2]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, USA

[3]Institute for Computational and Data Sciences, Pennsylvania State University, USA

Email for all authors: sweetalana2000@gmail.com, shirinnataneli@gmail.com, shengmia@usc.edu, jazlyn.mooney@gmail.com, szpiech@psu.edu

*Author for correspondence: Zachary A Szpiech, Jazlyn A Mooney

Mailing addresses with zip codes for all authors:

[1]514A Wartik Laboratory, University Park, PA, USA 16802

[2]1050 Childs Way, Ray R. Irani Hall, Los Angeles, CA, USA, 90089

ORCID:
Sweetalana: 0009-0000-5450-5934
Zachary A Szpiech: 0000-0001-6372-8224
Jazlyn A Mooney: 0000-0002-2369-0855
Shirin Nataneli: 0000-0001-7880-9967
Shengmiao Huang: 0009-0000-8048-9683

Keywords: runs of homozygosity, dogs, inbreeding, GWAS, complex traits

Abstract

Runs of homozygosity (ROH) are genomic regions that arise when two copies of identical haplotypes are inherited from a shared common ancestor. In this study, we leverage ROH to identify associations between genetic diversity and non-disease phenotypes in *Canis lupus familiaris* (dogs). We find significant association between the ROH inbreeding coefficient (F_{ROH}) and several phenotypic traits. These traits include height, weight, lifespan, muscled, white coloring of the head and chest, furnishings, and fur length. After correcting for population structure, we identified more than 45 genes across the examined quantitative traits that exceed the threshold for suggestive significance. We observe distinct distributions of inbreeding and elevated levels of long ROH in modern breed dogs compared to more ancient breeds, which aligns with breeding practices during Victorian era breed establishment. Our results highlight the impact of non-additive variation and of polygenicity on complex quantitative phenotypes in dogs due to domestication and the breed formation bottleneck.

Introduction

For over a century, scholars and dog-enthusiasts alike have sought to unravel the complex evolutionary history of man's best friend (Darwin, 1868). *Canis lupus familiaris* (dogs) have intrigued researchers, across scientific fields, due to their close knit ties to our species (Nagasawa et al., 2015) and unique genetic features (Boyko, 2011; Freedman et al., 2014; Perri et al., 2021). While the exact origins of domestication remain elusive, our growing understanding of this process offers new insight into the more recent evolutionary history of the species.

In the 19th century, Darwin suggested that the broad range of features exhibited in modern breeds is conducive to a descendancy from multiple canid species, including wolves and jackals. In his accounts on variation, he argues that a single ancestral origin from the gray wolf (*Canis lupus*) seems to contradict the historical and archeological evidence available. Given that many wild species of the genus Canis have the potential to interbreed, Darwin's proposal held fast for many years (Darwin, 1868). Advances in genome technology provided scientists with a new approach to studying dog ancestry, offering insights that complement and, in some cases, challenge the conclusions drawn from archaeological and historical evidence (Gray, 1972). One commonly used approach is mitochondrial DNA (mtDNA) analysis which allows researchers to trace maternal ancestry. Using samples collected from dogs, wolves, and several other wild canid species, researchers found the similarity between dogs and gray wolf to be significantly higher than that of dogs and any other wild canid species (Vilà et al., 1997; Wayne, 1993). In recent years, whole genome sequence (WGS) data has increasingly focused on the dog-wolf relationship, further supporting this finding and making it a primary area of investigation in contemporary studies (Freedman et al., 2014; Meadows et al., 2023; Ostrander et al., 2019).

While most agree that dogs were the first domesticated species, establishing their spatial and temporal roots has been a challenge (Larson & Bradley, 2014). The earliest depictions of human-canine ties come by means of cave paintings discovered in Saudi Arabia (Guagnin et al., 2018). This cooperative hunting scene is believed to be 8000 to 9000 years old, but evidence gathered from burial sites predicted an earlier origin of domestication (Morey, 2006). The first dates back approximately 12,000 years ago to the Natufian in Northern Israel, where they discovered three sets of dog remains buried with a human (Davis & Valla, 1978; Tchernov & Valla, 1997). Additionally, a dog-like mandible was recovered at a burial site in Bonn-Oberkassel, Germany.

This Northern European sample is believed to be about 14,000 years old (Benecke, 1987; Nobis, 1979), however genetic work has added an additional layer of evidence for an even older wolf-dog divergence.

Genomic work, based on mtDNA and WGS data, currently offers two potential explanations for the origins of domestication. The first is that early dogs were a result of a single domestication event from gray wolves in East Asia between 16,000 and 100,000 years ago. The hypothesis stems from the limited haplotypic diversity in dogs, which may suggest a genetic bottleneck where domesticated dogs originated from a small-closely-related population of gray wolves (Vilà et al., 1997). Additionally, high levels of genetic diversity from the region south of the Yangtze River, indicates an older divergence than other areas in Europe and Asia. (Pang et al., 2009; Savolainen et al., 2002; G.-D. Wang et al., 2016).

The second genomic narrative centers upon the theory that haplogroups entered the dog lineage at different times and places. Using nuclear genomic single nucleotide polymorphism (SNP) analysis, one study found evidence for allele sharing in Asian dog breeds and Asian wolves as well as European dog breeds and European wolves (vonHoldt et al., 2010). This work suggests that there were several founder events derived from separate wolf populations. Further WGS work has documented significant genetic differences in Eastern and Western Eurasian dog populations, providing further support for multiple domestication events (Bergström et al., 2022; Fan et al., 2016).

While the early evolution of dogs from wolves marked a significant shift between species, the more recent history of established dog breeds has shaped evolution within the species itself. During the Victorian era, the selective breeding of aesthetically desirable traits led to the emergence of

91 hundreds of new breeds. Breed establishment typically involved a limited number of founding
92 members, resulting in high levels of inbreeding and dramatic loss of genetic diversity (Bannasch
93 et al., 2021; Boyko, 2011; Cecchi et al., 2013; Mastrangelo et al., 2018; Vaysse et al., 2011; Yordy
94 et al., 2020). Founder effects, such as this, restrict the gene pool and result in long stretches of
95 DNA sharing between individuals. Shared segments derived from a single common ancestor are
96 said to be identical-by-descent (IBD). Regions inherited IBD, commonly referred to as runs of
97 homozygosity (ROH) when found within an individual, reflect a simplified trait architecture in
98 dogs that facilitates the study of phenotypic associations between breeds (Boyko, 2011; Plassais
99 et al., 2019; Sams & Boyko, 2019).

100 Studies of ROH in humans have revealed associations with numerous complex disease and non-
101 disease phenotypes, which provide valuable insight on the impact of genetic diversity on health
102 and evolutionary processes (Bacolod et al., 2008; Clark et al., 2019; Johnson et al., 2018; Keller
103 et al., 2012; Moreno-Grau et al., 2021; Pemberton & Szpiech, 2018). Similar work has been done
104 in dogs, providing evidence of a relationship between ROH and several disease phenotypes
105 (Boyko, 2011; Mooney et al., 2021; Sams & Boyko, 2019). Additionally, previous studies have
106 explored the influence of genetic variants on quantitative traits such as; leg length, wrinkled skin,
107 coat color, hair length, and skeletal shape (Boyko et al., 2010; Cadieu et al., 2009; Candille et al.,
108 2007; Parker et al., 2009; Sutter et al., 2007). Despite these advances, associations with complex
109 non-disease phenotypes in dogs are relatively unexplored (Hayward et al., 2016).

110 In this study, we investigate the relationship between the genomic consequences of domestication
111 and complex (non-disease) trait architecture by using the distribution of ROH to provide insights
112 on complex phenotypic architecture. These associations highlight potential to discover non-
113 additive variation and phenotypes. To examine the relationship between ROH and non-disease

114 phenotypes in dogs, we analyzed 556 canid whole-genome sequences and characterized the
115 genomic distribution of ROH. After accounting for breed structure, we test for associations
116 between the total ROH coverage of the genome (as measured by F_{ROH}) among 13 breed groups
117 with 13 phenotypes. Finally, we perform an ROH-mapping genome-wide association study and
118 identify multiple genetic variants associated with phenotype, namely height, weight, and lifespan.

119 # Methods

120 ## Data filtering and categorizations

121 In this study, we utilized whole genome sequences from 722 canids, which included various wild
122 species, dingoes, and domestic dogs (Plassais et al., 2019). This data can be accessed via NCBI
123 accession number PRJNA448733. All filtering was accomplished using BCFtools (Danecek et al.,
124 2021) where we retained only biallelic SNPs and sites of genotype quality score greater or equal
125 to 20. Additionally, we removed any missing genotype rates above 10% and removed variants
126 from sex chromosomes. Our filtered data consisted of 4,053,761 biallelic loci for all 38 pairs of
127 autosomal chromosomes. Given the nature of our study, we removed all wild canids, mixed-breed
128 individuals, and samples for which breed information was unavailable. Each remaining sample
129 was categorized by common name or breed (Plassais et al., 2019). These classifications were
130 informed by neighbor joining trees of domestic dogs (vonHoldt et al., 2010) and American Kennel
131 Club (AKC) groupings (*Dog Breeds - Types Of Dogs*, n.d.). This resulted in 13 breed groups
132 namely, Ancient Spitz (31 individuals and 11 breeds), Herding (76 individuals and 15 breeds),
133 Mastiff-like (65 individuals and 15 breeds), Retrievers (50 individuals and 6 breeds), Scent Hound
134 (23 individuals and 10 breeds), Small Terrier (102 individuals and 8 breeds), Terriers (16
135 individuals and 7 breeds), Spaniels (22 individuals and 8 breeds), Toy Dogs (18 individuals and 7

136 breeds), Working Dogs (44 individuals and 10 breeds), Sight Hound (21 individuals and 9 breeds),
137 Village Dogs (75 individuals and 13 regions), and Chinese Indigenous dogs (15 individuals).
138 Overall, our working dataset included 558 individuals and 106 different breeds (Table S1).

139 **Calling Runs of Homozygosity**

140 To call runs of homozygosity (ROH) we used a likelihood-based inference method called GARLIC
141 v1.6.0a (Szpiech et al., 2017). This method uses a logarithm of odds (LOD) score measure of
142 autozygosity, which is applied in a sliding window for the entire genome (Pemberton et al., 2012).
143 GARLIC requires input data with genotype and sample information in the form of TPED and
144 TFAM files, which were generated using PLINK 2.0 (C. C. Chang et al., 2015). We generated a
145 TGLS file to obtain per-genotype likelihoods. We employed genotype likelihood data in the form
146 of GQ to account for errors in phred-scaled probability. Next, a window size of 100 SNPs was
147 chosen based on SNP density, with the window incrementally advancing by 10 SNPs at each step.
148 GARLIC offers built-in ROH length classification, for which we defined 6 classes: <1 Mbp, 1-2
149 Mbp, 2-3 Mbp, 3-4 Mbp, 4-5 Mbp, and >5 Mbp. To account for sample size differences in allele
150 frequency estimation, we set the number of resamples to 40. All other flags were set to default,
151 and the following parameters were used: --auto-winsize --auto-overlap-frac --winsize 100 --
152 centromere --size-bounds 1000000 2000000 3000000 4000000 5000000 --tped --tfam --tgls –gl-
153 type GQ --resample 40 --out. To eliminate segments that were noisy, very short, and common, we
154 filtered to retain ROH segments longer than 0.5 Mbp.

155 We conducted independent ROH calling for breeds with at least 10 individuals. Breeds consisting
156 of fewer than 10 members were integrated into their respective breed groups (Table S2).
157 Subsequently, ROH calls from all breeds, irrespective of individual numbers, were combined into

158 their respective breed groups for further analysis. For each individual, we examined: 1) the
159 relationship between coverage and F_{ROH} (Figure S1) and 2) the relationship between the total
160 number of runs of homozygosity (n_{ROH}) and total length of runs of homozygosity (s_{ROH}). For each
161 individual and breed group, we calculated the arithmetic mean and range of both total n_{ROH} and
162 total s_{ROH}, while binning them based on ROH size classification. The longest total s_{ROH} (2141.22
163 Mbp) and the highest total n_{ROH} (1039) were identified in two individuals - PER00747 and
164 PER00393, respectively - from the small terrier Yorkshire breed. PER00747 was sequenced at a
165 low depth (~2x), which inflated s_{ROH}. To mitigate the effects of low depth PER00747 and
166 PER00393 were removed from downstream analysis.

167 **Inbreeding coefficient**

168 For each sample and breed group, we computed the inbreeding coefficient, F_{ROH}, defined as the
169 fraction of autosomal genome in ROH regions (McQuillan, 2009):

170 $$F_{ROH} = \frac{s_{ROH}}{L_{Auto}}$$

171 Where s_{ROH} is the sum of lengths of runs of homozygosity in an individual's genome and L_{Auto} is
172 the length of autosomal genome in base pairs. Autosomal genome length of 2,203,764,842 base
173 pairs was used since variants were called with the CanFam3.1 (NCBI RefSeq assembly:
174 GCF_000002285.3) reference genome.

175 **ROH Sharing Matrix**

176 We generated an ROH sharing matrix, a square matrix that quantifies the sum of ROH overlaps
177 between two individuals. This allowed us to analyze shared ROH patterns among dog breeds. To

178 populate this matrix, we first generated individual ROH bed files for each sample. We then used
179 BEDTools (Quinlan & Hall, 2010) to identify overlapping regions between pairs of ROH data.
180 The following parameters were used: bedtools intersect -wao.

181 Phenotype data

182 We updated the phenotypic data (Plassais et al., 2019) for 13 traits: bulky, drop ears, furnish,
183 hairless, height, large ears, length of fur, lifespan, long legs, muscled, weight, white chest, and
184 white head. For the three continuous traits, namely height, weight, and lifespan, we used average
185 values obtained from AKC (*Dog Breeds - Types Of Dogs*, n.d.) (Table S3). We chose to use breed
186 average values because this was previously shown to be a reliable measure for association tests
187 (Boyko et al., 2010; Hayward et al., 2016; Jones et al., 2008; Plassais et al., 2017, 2019; Rimbault
188 et al., 2013). When sex-specific information was available, it was incorporated into the dataset;
189 otherwise, we applied the same values for both males and females. We applied min-max scaling
190 to normalize the three continuous traits and further categorized them into small and large groups
191 based on the average weight (21.66kg) of all individuals (Table S4, Table S5, Table S6). For the
192 remaining 10 traits, we averaged the values by the phenotypic data available for their respective
193 breeds (Table S7). We used binary encoding to assign a value of 1 to individuals expressing the
194 trait and 0 to those who did not.

195 Association tests

196 We computed the association between F_{ROH} and the 13 phenotypes using GMMAT, an R package
197 that performs association tests with generalized linear mixed models (GLMM) (Chen et al., 2016).
198 To fit the GLMM, we used the built-in function glmmkin, which allowed us to examine the

199 continuous traits (height, weight, and lifespan) as the quantitative phenotype traits. This also
200 enabled us to include individual F_{ROH} as a covariate and use the ROH sharing matrix as a kinship
201 matrix. We fit the model assuming a Gaussian distribution for the continuous phenotypes and used
202 the identity link function. For associations between breed average height, weight, and lifespan with
203 F_{ROH}, we used the updated average phenotypic values of the breed for all 466 individuals. Based
204 on our association tests and previous studies (Fareed & Afzal, 2014; Kim et al., 2018), weight and
205 F_{ROH} have been observed to have a significant association, thus we included them as covariates in
206 our models. Weight and the interaction of weight and F_{ROH} were used as covariates across three
207 data subsets (all individuals, small individuals, and large individuals) (Table S8). To fit the GLMM
208 to the 10 binary traits, we specified the binomial distribution as the family and used the logit link
209 function. Samples with no information available on the presence or absence of these phenotypes
210 were removed from analysis. Due to lack of phenotype information, Chinese indigenous dogs and
211 village dogs were excluded from the association tests.

212 **ROH-mapping GWAS**

213 To explore the biological mechanisms through which ROH-associated SNPs influence phenotypic
214 traits, we conducted Genome-Wide Association Studies (GWAS). We generated a dataset that
215 examined the occurrence of SNPs in ROH regions using BCFtools intersect and BCFtools subtract
216 (Danecek et al., 2021). Any SNPs located outside of ROH regions were removed. Using GMMAT,
217 we fit GLMM models assuming a recessive genetic model for quantitative traits (height, weight,
218 and lifespan), incorporating ROH-associated SNPs as covariates, an ROH sharing matrix for
219 kinship structure, and a Gaussian distribution with an identity link function (Chen et al., 2016).
220 Additionally, given that breed structure has a particularly strong effect in dogs, we utilized a

221 kinship matrix quantified by pairwise ROH sharing between individuals when associating traits
222 with F_{ROH}. This procedure follows from an approach from a study on disease traits (Mooney et al.,
223 2021). Lastly, we fit GLMM models using weight and the product of weight and F_{ROH} as covariates
224 for the three data subsets (all individuals, small individuals, and large individuals) (Table S9). To
225 obtain the effective sample size, adjusted for auto correlations between sites, we used the option
226 effectiveSize within the coda R package (Plummer et al., 2006). To calculate the number of
227 independent tests we used p-values for every SNP as the input, thus incorporating autocorrelation
228 caused by LD. This approach was previously shown to be effective for correcting for structure
229 (Shriner et al., 2011). To obtain the Genome-Wide Significance (GWS) and Suggestive-Wide
230 Significance (SWS) thresholds, we used the Bonferroni correction with respect to effective sample
231 size (Li et al., 2012; W. Yang et al., 2014; Yuan et al., 2020):

232 $$\text{Genome-wide significance threshold} = -log10\left(\frac{0.05}{Effective\ Sample\ Size}\right)$$

233 $$\text{Suggestive-wide significance threshold} = -log10\left(\frac{0.1}{Effective\ Sample\ Size}\right)$$

234 To correct for population stratification, we calculated the genomic inflation factor, λ (Sofer et al.,
235 2021; Van Den Berg et al., 2019). We corrected the log-scaled p-values by dividing them by the
236 genomic inflation factor (Table S9). To visualize the GWAS summary statistics, we used an R
237 package, qqman to create Q-Q and Manhattan plots (Turner, 2014).

238 # Results

239 **ROH distribution**

We used GARLIC to call ROH per individual and divide ROH into our pre-specified length classes. This allowed us to explore the variance per-population within each class. ROH were categorized into 6 length classes: Class A (0.5 - 1 Mbp), Class B (1 - 2 Mbp), Class C (2 - 3 Mbp), Class D (3 - 4 Mbp), Class E (4 - 5 Mbp), and Class F (>5 Mbp). For our 556 samples, the mean number of ROH (n_{ROH}) was 37.06 and the mean length of ROH (s_{ROH}) was 94.17 Mbp. These parameters were also computed for each ROH length class: Class A (n_{ROH} = 99.44; s_{ROH} = 68.39 Mbp), Class B (n_{ROH} = 50.29; s_{ROH} = 70.36 Mbp), Class C (n_{ROH} = 21.60; s_{ROH} = 52.81 Mbp), Class D (n_{ROH} = 13.07; s_{ROH} = 45.23 Mbp), Class E (n_{ROH} = 9.21; s_{ROH} =41.15 Mbp), and Class F (n_{ROH} = 28.77; s_{ROH} = 287.05 Mbp) (Table 1, Table 2).

249 ***TABLE 1.*** *nROH across all breed groups and ROH length classification. The table represents*
250 *the range and mean ± standard deviation of the number of runs of homozygosity for each breed*
251 *group and ROH length classification.*

Breed Group	Mean n_{ROH}	Range n_{ROH}	Mean n_{ROH} for Class A (0.5-1 Mbp)	Mean n_{ROH} for Class B (1 - 2 Mbp)	Mean n_{ROH} for Class C (2 - 3 Mbp)	Mean n_{ROH} for Class D (3 - 4 Mbp)	Mean n_{ROH} for Class E (4 - 5 Mbp)	Mean n_{ROH} for Class F (> 5 Mbp)
Ancient Spitz	47.10 ± 58.20	3 - 338	133.42 ± 85.63	65.13 ± 40.00	24.71 ± 11.23	13.90 ± 4.31	9.87 ± 3.97	35.58 ± 18.29
Retrievers	38.83 ± 34.58	0 - 138	106.42 ± 13.75	53.96 ± 9.05	23.98 ± 6.39	13.1 ± 4.31	9.94 ± 3.32	26.3 ± 11.77
Mastiff Like	38.37 ± 26.65	0 - 117	78.6 ± 14.12	54 ± 17.53	27.46 ± 11.19	16.91 ± 7.55	12.18 ± 6.53	41.08 ± 19.83
Herding	37.88 ± 33.97	0 - 257	97.76 ± 32.29	48.47 ± 10.17	21 ± 7.22	13.91 ± 5.79	10.8 ± 5.51	35.32 ± 19.08
Scent Hound	40.25 ± 37.52	1 - 160	113.26 ± 21.76	53.65 ± 9.85	22.7 ± 6.35	13.3 ± 4.87	9.09 ± 4.92	27.48 ± 12.44
Small Terrier	40.93 ± 41.19	0 - 277	113.58 ± 39.05	60.27 ± 17.74	24.08 ± 9.18	14.31 ± 6.32	8.98 ± 4.61	24.38 ± 16
Terriers	30.56 ± 33.01	9 - 122	90.75 ± 15.75	56.25 ± 11.22	31.81 ± 7.75	18.88 ± 6.21	15.88 ± 4.66	53.75 ± 10.77
Spaniels	36.29 ± 35.23	5 - 144	101.18 ± 20.19	61 ± 9.23	28.73 ± 5.37	18.14 ± 5.35	14 ± 5.19	41.64 ± 14.62
Sight Hound	45.29 ± 48.55	0 - 192	138.67 ± 31.05	63.33 ± 20.63	21.67 ± 8.98	13.86 ± 7.57	8.19 ± 4.76	26 ± 15.86
Toy Dogs	41.19 ± 40.77	1 - 187	114.56 ± 31.06	56.17 ± 25.99	24.06 ± 13.48	14.78 ± 9.78	9.44 ± 6.46	28.11 ± 15.34
Working Dogs	38.29 ± 27.43	1 - 129	85.18 ± 16.07	50.02 ± 15.37	27.36 ± 12.12	17 ± 7.86	11.82 ± 5.51	38.34 ± 13.11
Village Dogs	20.5 ± 31.83	0 - 261	80.12 ± 33.69	23.79 ± 13.69	6.33 ± 5.77	3.51 ± 4.76	2.09 ± 2.64	7.15 ± 11.03
Chinese Indigenous Dogs	14.36 ± 22.91	0 - 103	61.53 ± 16.63	14.67 ± 5.79	4 ± 3.31	1.27 ± 1.29	1 ± 1.32	3.67 ± 5.37

252

253 ***TABLE 2.*** *sROH across all breed groups and ROH length classification. The table represents*
254 *the range and mean ± standard deviation of the sum of runs of homozygosity for each breed*
255 *group and ROH length classification.*

Breed Group	Mean s_{ROH} (Mbp)	Range s_{ROH} (Mbp)	Mean s_{ROH} for Class A (0.5-1 Mbp)	Mean s_{ROH} for Class B (1-2 Mbp)	Mean s_{ROH} for Class C (2-3 Mbp)	Mean s_{ROH} for Class D (3-4 Mbp)	Mean s_{ROH} for Class E (4-5 Mbp)	Mean s_{ROH} for Class F (> 5 Mbp)
Ancient Spitz	119.56 ± 162.09	13.73 - 1113.00	92.29 ± 59.41	90.67 ± 54.30	60.33 ± 27.20	48.30 ±	44.14 ± 17.51	381.60 ± 255.7
Retrievers	89.28 ± 86.63	0 - 598.02	72.96 ± 9.44	75.50 ± 13.25	57.37 ± 16.00	45.22 ± 15.07	44.47 ± 14.90	240.17 ± 125.98
Mastiff Like	126.47 ± 176.91	0 - 1029.28	54.77 ± 9.71	76.97 ± 25.95	67.21 ± 27.55	58.57 ± 25.92	54.38 ± 29.38	446.89 ± 247.25
Herding	105.25 ± 139.34	0 - 861.74	67.24 ± 21.71	67.53 ± 14.78	51.27 ± 17.64	48.23 ± 20.12	48.26 ± 24.73	348.94 ± 206.94
Scent Hound	99.19 ± 113.14	4.59 - 946.54	77.78 ± 14.73	74.84 ± 15.22	55.26 ± 15.67	45.90 ± 16.82	40.62 ± 21.78	300.75 ± 159.58
Small Terrier	89.41 ± 96.13	0 - 690.33	78.20 ± 26.38	84.24 ± 24.77	59.07 ± 22.76	49.63 ± 22.10	40.18 ± 20.51	225.15 ± 170.89
Terriers	75.86 ± 116.57	28.77 - 951.29	63.93 ± 10.90	79.69 ± 16.00	77.98 ± 19.72	65.84 ± 21.84	71.06 ± 20.79	541.71 ± 161.42
Spaniels	94.03 ± 127.54	23.23 - 856.91	70.13 ± 12.83	85.21 ± 12.79	70.39 ± 13.56	62.52 ± 18.65	62.55 ± 23.16	427.24 ± 190.19
Sight Hound	94.65 ± 96.81	0 - 591.60	95.24 ± 21.50	86.87 ± 29.21	53.04 ± 22.07	47.98 ± 26.13	36.54 ± 21.32	248.26 ± 149.49
Toy Dogs	97.01 ± 107.57	4.12 - 602.43	78.92 ± 22.17	79.01 ± 36.61	58.31 ± 33.13	50.45 ± 33.72	42.33 ± 28.82	273.02 ± 162.01
Working Dogs	112.89 ± 130.13	4.74 - 658.42	58.67 ± 11.14	70.67 ± 22.08	66.54 ± 29.19	58.83 ± 27.51	52.76 ± 24.70	369.85 ± 139.08
Village Dogs	33.24 ± 59.75	0 - 554.84	54.07 ± 23.09	32.65 ± 19.59	15.30 ± 13.97	12.04 ± 16.34	9.35 ± 11.72	76.00 ± 127.67
Chinese Indigenous Dogs	19.82 ± 35.06	0 - 309.27	41.03 ± 11.34	20.00 ± 7.64	9.53 ± 7.68	4.27 ± 4.24	4.41 ± 5.87	39.67 ± 75.16

256

257 Domestic dog clades showed a distinct pattern of genetic diversity, with higher levels of long ROH
258 (>5 Mbp) when compared with Chinese indigenous dogs (light green) and village dogs (royal blue)
259 (Figure 1 and Figure 2). The highest total n_{ROH} (641) was observed in Dingo01 and the highest
260 total s_{ROH} (1643.98 Mbp) was observed in NGSD2, both individuals were from the Ancient Spitz
261 breed group (Table S10). Terriers had the smallest mean n_{ROH} and the smallest mean s_{ROH} when
262 compared with the remaining 11 breed groups (Table 1, Table 2).

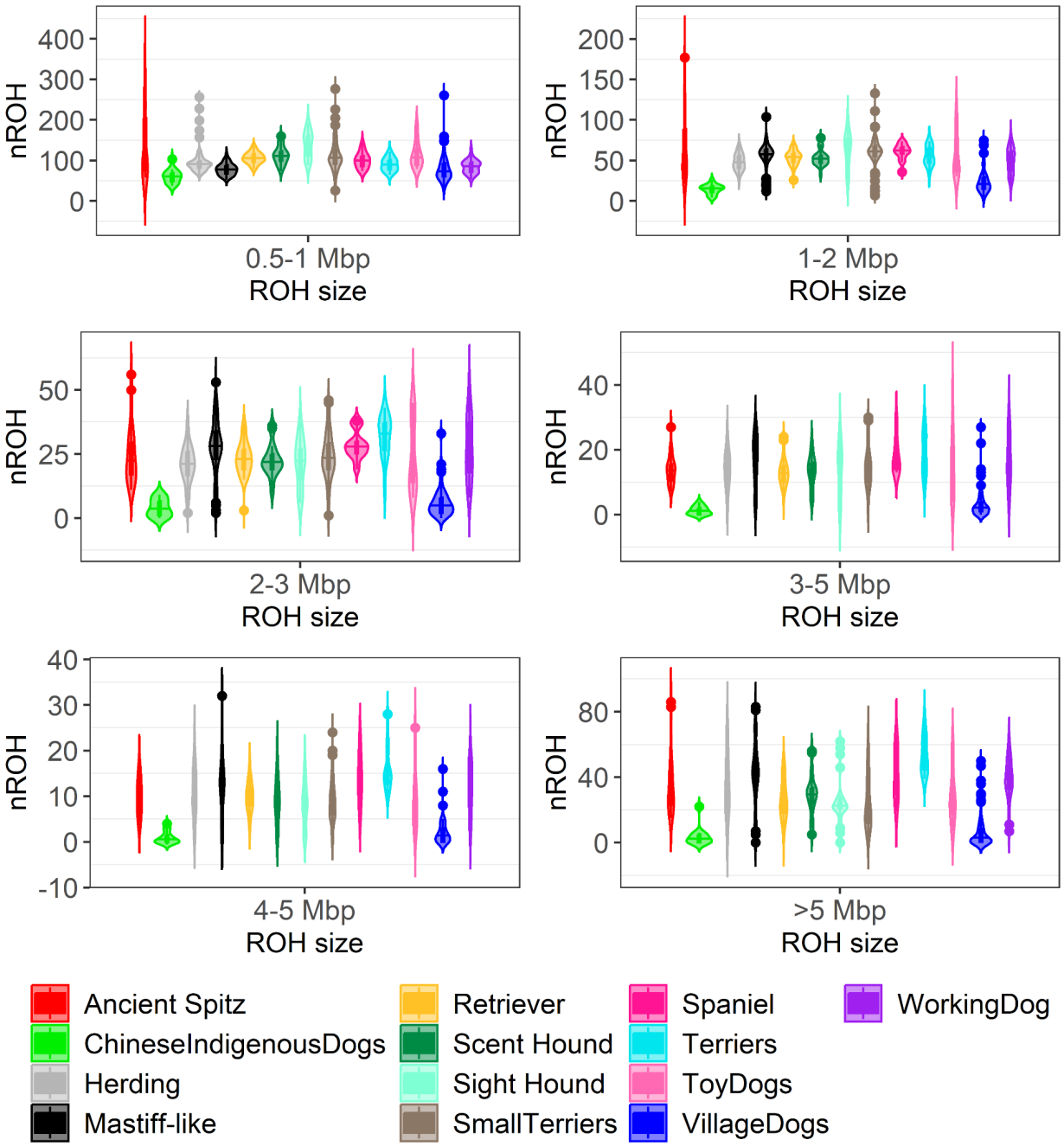


FIGURE 1. *Relationship between number of runs of homozygosity (nROH) and ROH size class in Mbp for all breed groups. The x-axis represents the ROH size in Mbp. The y-axis represents the number of ROH. Each violin plot depicts the distribution of the total number of ROH within length ranges 0.5-1 Mbp, 1-2 Mbp, 2-3 Mbp, 3-4 Mbp, 4-5 Mbp and >5 Mbp for each breed group.*

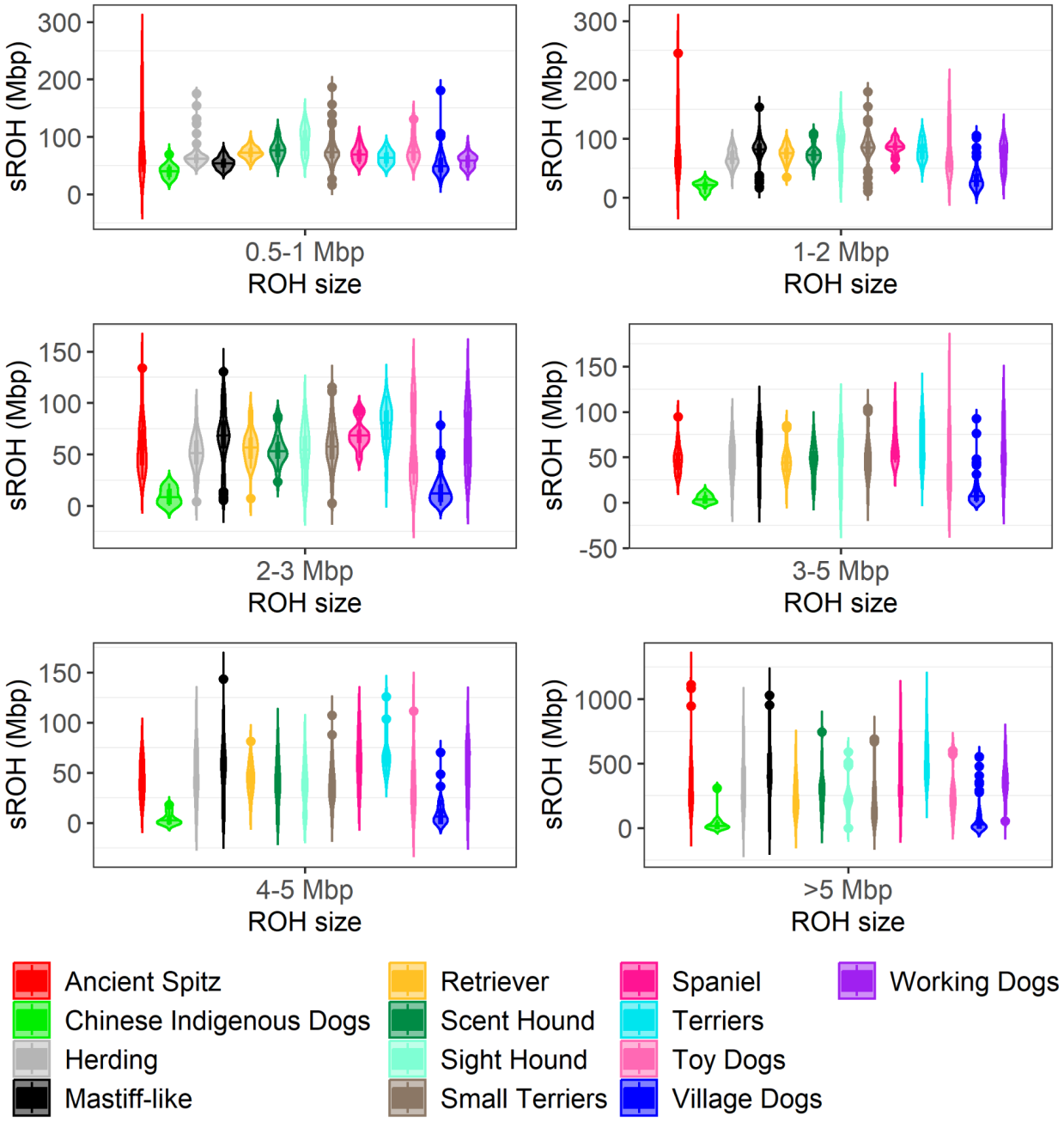


FIGURE 2. *Relationship between sum of runs of homozygosity (sROH) and ROH size class in Mbp. The x-axis represents the ROH size in Mbp. The y-axis represents the sum of runs of homozygosity. Each violin plot depicts the distribution of the sum of ROH within length ranges 0.5-1 Mbp, 1-2 Mbp, 2-3 Mbp, 3-4 Mbp, 4-5 Mbp and >5 Mbp for each breed group.*

Next, we explored the relationship between the total number of ROH (n_{ROH}) and the total lengths of ROH (s_{ROH}) (Figure 3). We observed significantly higher mean values of both n_{ROH} and s_{ROH} for the domesticated breed dogs relative to the Chinese indigenous dogs and village dogs.

278 Breed dogs also tended to fall off of the y=x line when n_{ROH} was compared to s_{ROH}. The
279 accumulation of long ROH within breed dogs in the recent time was most drastic in the Terrier
280 clade, where we also observed the largest F_{ROH} (Table 3). We saw the largest variance in nROH
281 in the Small Terriers, which include breeds such as the Australian Terrier, Cairn Terrier, Jack
282 Russell Terrier, Norwich Terrier, Scottish Terrier, Tibetan Terrier, West Highland, White Terrier,
283 and Yorkshire Terrier. We also observed larger variance in n_{ROH} within the Ancient Spitz, Village
284 dogs, Herding, Mastiff-like, and Toy dogs. Conversely, Chinese Indigenous dogs, Terriers,
285 Spaniels, Retrievers, and Scent Hound clustered together quite tightly. The Spaniels fall
286 somewhere between the other clades.

287 Given the large variance in n_{ROH} we wondered whether we might be able to identify ROH
288 hotspots that were shared across clades or between domesticated and non-domesticated species.
289 We examined ROH coverage per site for each chromosome and focused on sites where more than
290 50% of individuals within a clade had a ROH (Figures S2-S17). There were clear hotspots within
291 clades on every chromosome. Some of these hotspots corresponded with recombination rate, such
292 as on chromosome 22, where we observed a low recombination rate in regions where there were
293 shared ROH hotspots among all individuals. Conversely, on chromosome 12 we observed a high
294 recombination rate, and very few shared ROH hotspots among individuals. Interestingly, we
295 observed very few shared ROH hotspots across different clades. One stand out chromosome was
296 28, where we observed a shared ROH hotspot among Ancient Spitz, Mastiff Like, Retrievers and
297 Terriers (Figure S14). Likewise on chromosome 32, we observed a shared ROH hotspot among
298 Mastiff-Like, Scent Hound, Terriers and Toy Dogs (Figure S15). The majority of hotspots were
299 shared within a clade. This was best shown on chromosome 17, where we found multiple clade-
300 specific ROH hotspots (Figure S10). The same pattern was seen on chromosomes 31 through 38,

alongside a few instances of sharing between Mastiff-like breeds and Terriers (Figures S14 - S17). Lastly, we rarely observed any shared peaks with village dogs and Chinese Indigenous dogs, which had relative few peaks in comparison to clades composed of breed dogs (Figures S4 - S17).

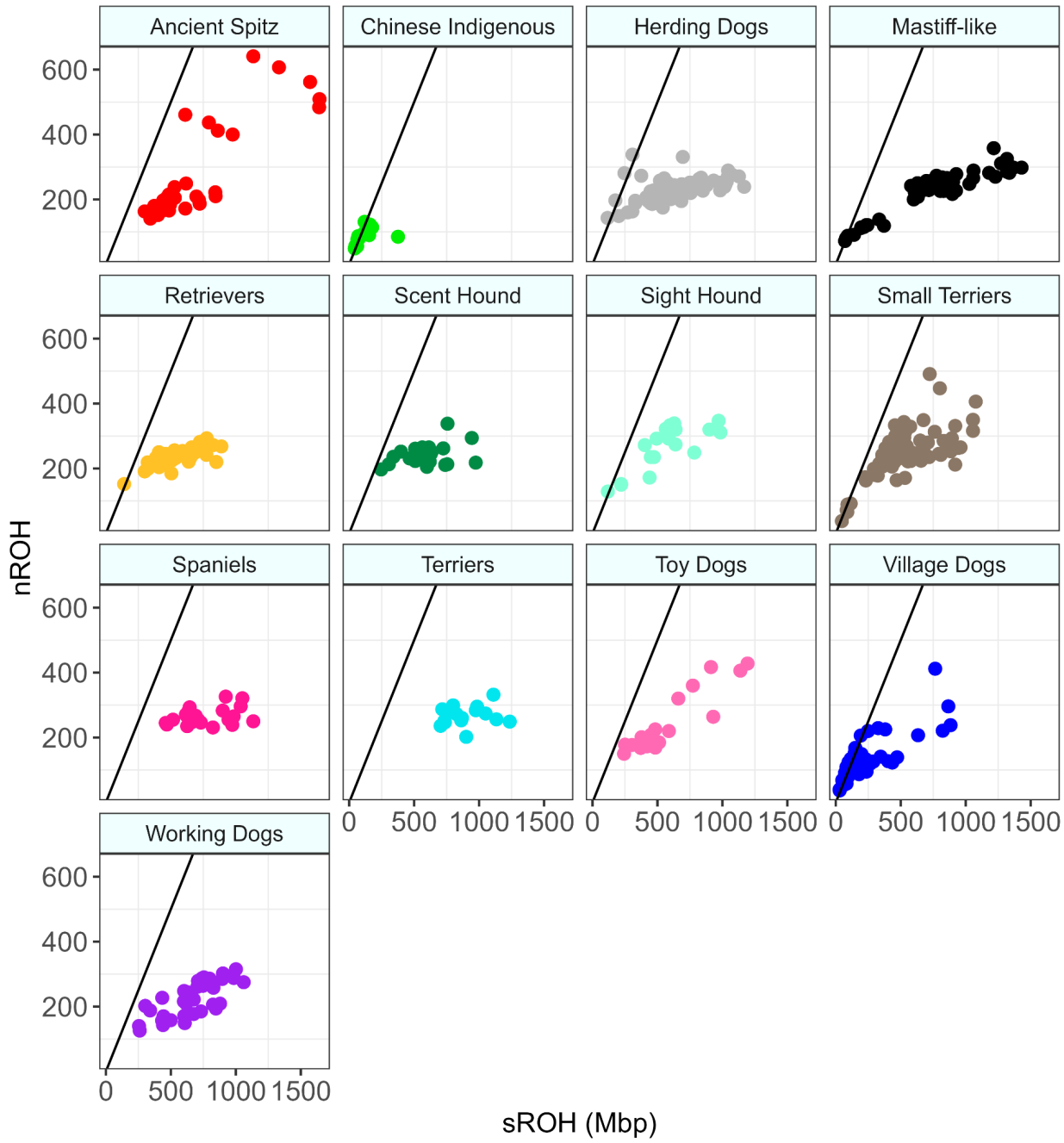


FIGURE 3. *Relationship between sum of runs of homozygosity (sROH) and number of runs of homozygosity (nROH) for each breed group. The x-axis and the y-axis represents the sROH and nROH, respectively. The black line is y=x.*

Next, we examined the level of inbreeding within clades by computing the inbreeding coefficient, F_{ROH}, for each individual and each breed group (Table 3, Table S11). To account for the distribution of F_{ROH} across all breed groups, we computed mean F_{ROH} (Table 3). High levels

313 of mean F_{ROH} were observed in most breed dogs clades. The mean F_{ROH} values in breed dogs
314 ranged from 0.24 ± 0.07 in Retrievers to 0.41 ± 0.07 in Terriers. The largest amounts of variance
315 were observed in the Mastiff-like and Ancient Spitz clades, which contain some of the earliest
316 breeds dog, such as the Tibetan Mastiff, Basenji and Shar-pei. The lowest mean values were
317 observed in non-breed dogs, namely Chinese Indigenous Dogs (0.05 ± 0.04) and Village Dogs
318 (0.09 ± 0.08).

319 ***TABLE 3.*** *F_{ROH} across all breed groups. The table represents the mean ± standard deviation of*
320 *F_{ROH} for each breed group*

Breed Group	Mean F_{ROH}
Terriers	0.41 ± 0.07
Spaniels	0.35 ± 0.09
Mastiff-Like	0.34 ± 0.15
Ancient Spitz	0.33 ± 0.17
Working Dog	0.31 ± 0.09
Herding	0.29 ± 0.11
Scent Hound	0.27 ± 0.08
Toy Dogs	0.26 ± 0.13
Sight Hound	0.26 ± 0.10
Small Terrier	0.24 ± 0.10
Retriever	0.24 ± 0.07
Village Dogs	0.09 ± 0.08
Chinese Indigenous Dogs	0.05 ± 0.04

321

322

Association between F_{ROH} and phenotypic traits

We were also interested in testing whether ROH could be used to detect non-additivity in quantitative (non-disease) phenotypes in dogs. Following an approach that was recently used in disease phenotypes in dogs (Mooney et al., 2021) and quantitative and disease phenotypes in humans (Clark et al., 2019; Malawsky et al., 2023), we searched for non-additive effects by correlating F_{ROH} and 13 breed averaged phenotypes (bulky, drop ears, furnish, hairless, height, large ears, length of fur, lifespan, long legs, muscled, weight, white chest, and white head). Three breed averaged quantitative phenotypes had significant associations with F_{ROH}: height ($\beta = 0.31$ and $p = 4.05 \times 10^{-3}$), weight ($\beta = 0.19$, $p = 1.19 \times 10^{-2}$), and lifespan ($\beta = -0.14$ and $p = 4.44 \times 10^{-2}$) (Figure 4). We found that as F_{ROH} increased by 1%; breed average height increased by 0.192 cm, breed average weight increased by 0.14 kg, and breed average lifespan decreased by 0.014 years. The statistical analysis, including normalized coefficient effect size (either β or log odds), p-values, and confidence intervals for quantitative traits (breed average height, breed average weight, and breed average lifespan) are shown in Table S8. Significant associations with F_{ROH} were present in 5 out of the 10 remaining qualitative phenotypes: muscled ($\beta = 7.96$ and $p = 9.68 \times 10^{-11}$), White chest ($\beta = 3.49$ and $p = 5.36 \times 10^{-5}$), White head ($\beta = 4.26$ and $p = 1.55 \times 10^{-4}$), length of fur ($\beta = -1.88$ and $p = 1.27 \times 10^{-2}$), and furnish ($\beta = -2.50$ and $p = 8.24 \times 10^{-3}$) (Figure 5, Table S12). Across all breed groups, as F_{ROH} increased, white chest, white head and muscled phenotypes also increased, whereas length of fur and furnish phenotypes decreased.

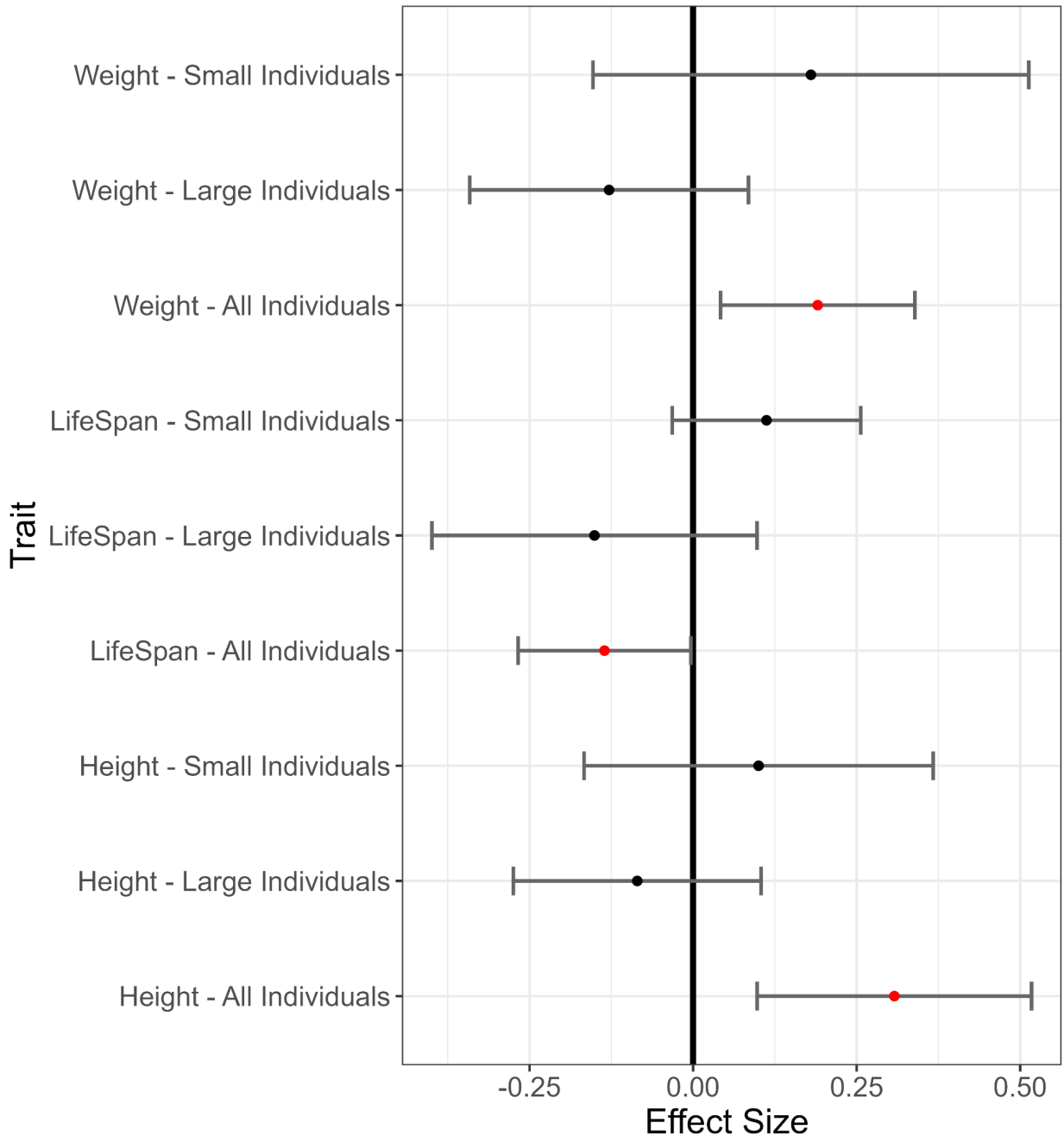


FIGURE 4. *Relationship between effect size and quantitative phenotypic traits. The y-axis represents the phenotypes for which the associations were tested. The x-axis represents the normalized beta coefficients (β) for their respective traits. A significant effect of FROH on a phenotype (nominal $p < 0.05$) is indicated with a red point. An effect size of greater than 0 indicates an increase of that trait with respect to F_{ROH}, and less than 0 represents the converse.*

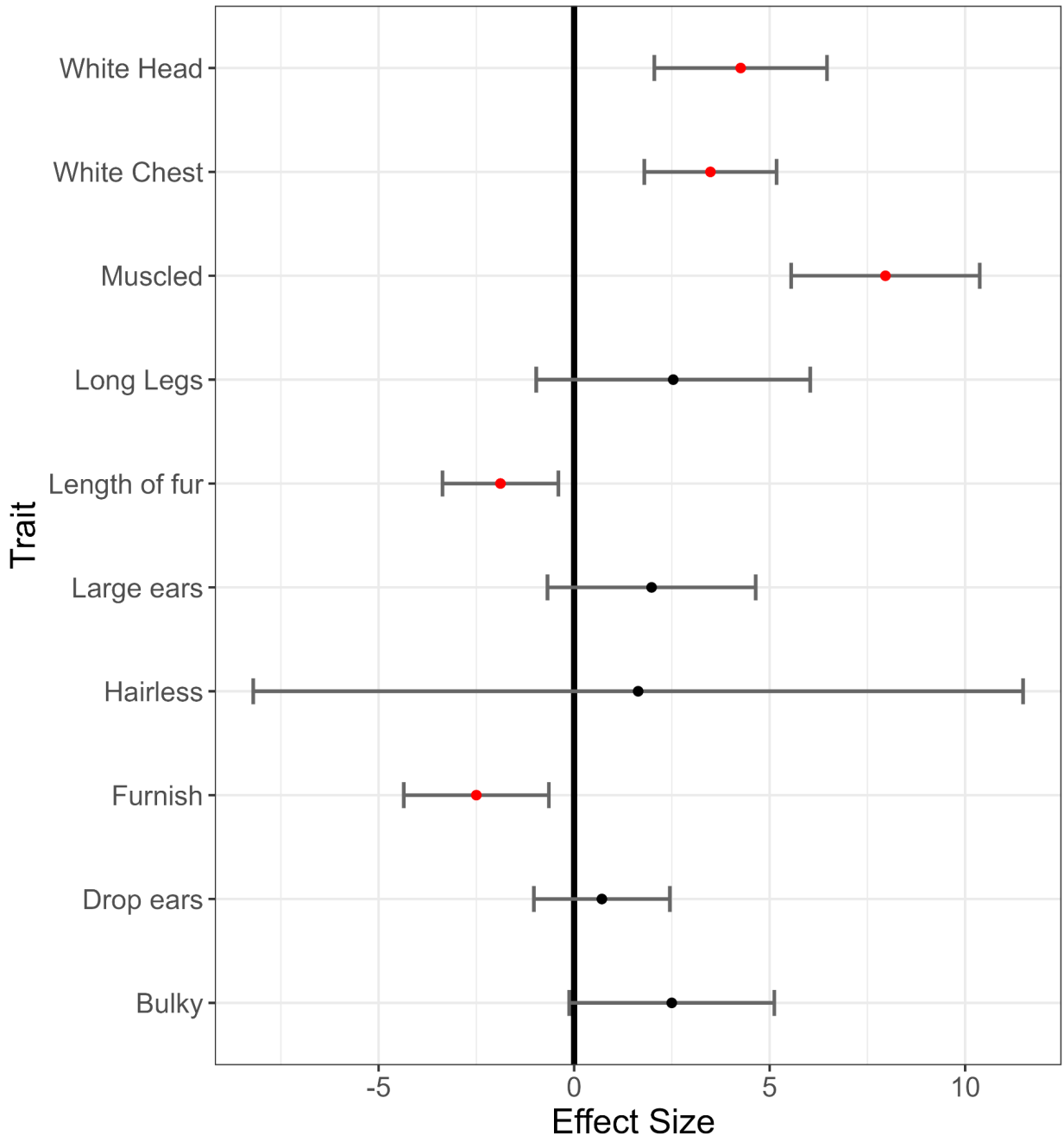


FIGURE 5. *Relationship between effect size and categorical phenotypic traits. The y-axis represents the phenotypes for which the associations were tested. The x-axis represents the log odds for their respective traits. A significant effect of* F_{ROH} *on a phenotype (nominal* $p < 0.05$*) is indicated with a red point. An effect size of greater than 0 indicates an increase of that trait with respect to* F_{ROH}*, and less than 0 represents the converse.*

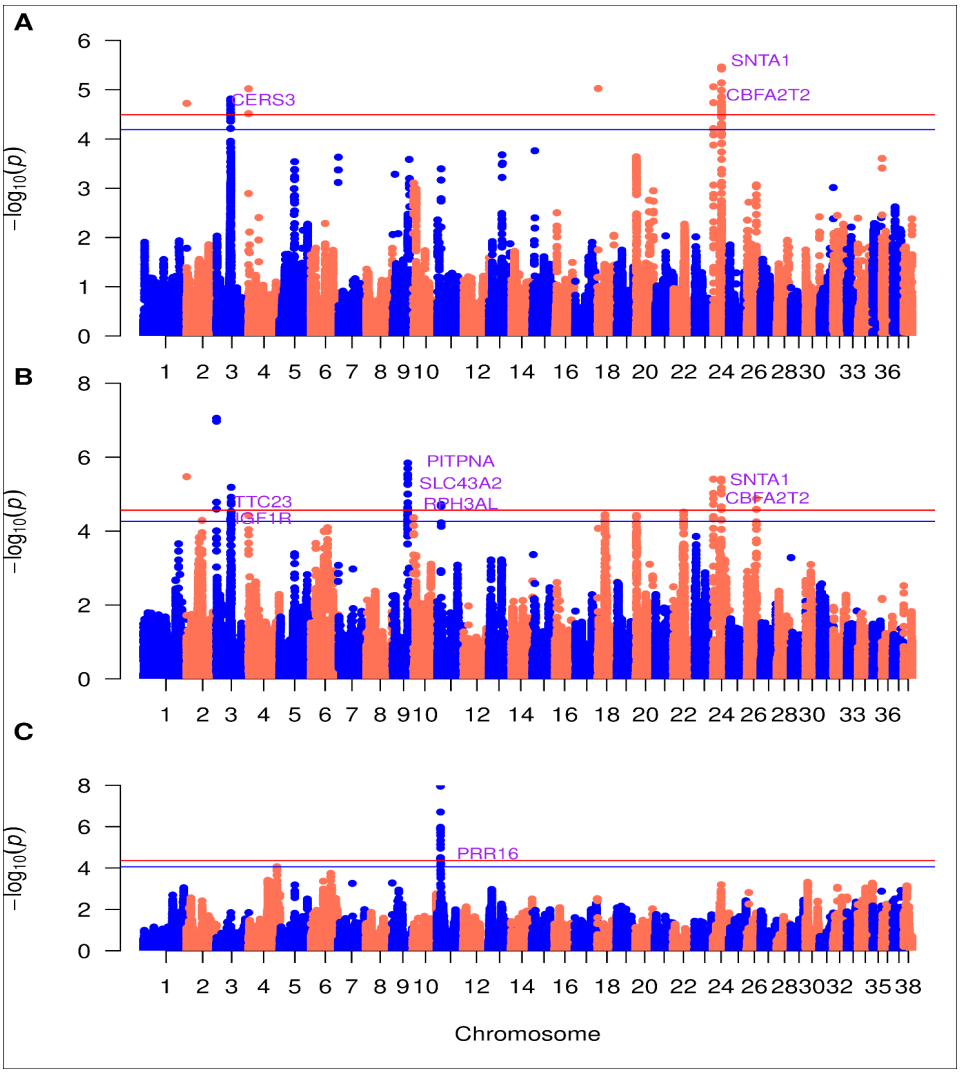
358 **ROH-mapping GWAS**

359 To identify the ROH-associated genomic regions that influence our quantitative traits of interest
360 (weight, lifespan, or height) we performed a GLMM-based GWAS (Figure 6, Figure 7, Figure 8,
361 and Figure S18-S26) using the presence or absence of ROH across individuals as the phenotype
362 predictor for each locus. For exact significant threshold values, see Tables S8 - S9.

363 For breed average height, we found 27 SNPs above the Suggestive-Wide Significance (SWS)
364 threshold and 18 SNPs above the Genome-Wide Significance (GWS) threshold when testing for
365 all individuals in our sample (Figure 6A). Notably, a single SNP on chromosome 24 (position
366 22647289) had the strongest associated genomic signal with height ($p = 3.46 \times 10^{-6}$). This
367 region corresponds with the gene *SNTA1*. Additionally, genes *CERS3* (7 SNPs) on chromosome
368 3, *CBFA2T2* (16 SNPs) on chromosome 24, and *SNTA1* (4 SNPs) on chromosome 24 were
369 observed to be above the SWS threshold and contained multiple SNPs within their transcription
370 regions (Table S13). SNPs linked to these three genes were also above the GWS threshold.

371 For small individuals, 34 SNPs were observed to be above the SWS threshold and 17 SNPs were
372 observed to be above the GWS threshold when tested for association with height (Figure 6B). A
373 single SNP on chromosome 9 (position 45711401) had the strongest associated genomic signal
374 ($p = 1.44 \times 10^{-6}$), which corresponds with the gene *PITPNA*. Across our small subgroup we
375 identified additional genomic variants above the SWS threshold for weight: on chromosome 3,
376 *CERS3* (3 SNPs), *IGF1R* (1 SNP), *LRRC28* (1 SNP), and *TTC23* (3 SNPs); on chromosome 9,
377 *PITPNA* (4 SNPs), *SLC43A2* (2 SNPs), *VPS53* (2 SNPs), and *RPH3AL* (4 SNPs); on
378 chromosome 18, *HGF* (2 SNPs); on chromosome 22, *COMMD6* (1 SNP); on chromosome 24,
379 *CBFA2T2* (7 SNPs) and *SNTA1* (4 SNPs) (Table S13). Only SNPs associated with *TTC23*,

380 *PITPNA*, *SLC43A2*, *RPH3AL*, CBFA2T2, and SNTA1 were above the GWS threshold. For large
381 individuals, a single SNP on chromosome 11 (position 9943557), which is linked to *PRR16*, was
382 observed to be above the SWS threshold for height (Figure 6C). For an extended table containing
383 all of the observed genes associated with height refer to Table S13.



384

385 ***FIGURE 6.*** *Manhattan plots for GLMM-based GWAS using presence/absence of ROH for body*
386 *height. A. All individuals B. Small Individuals C. Large Individuals. The x-axis represents the*
387 *genomic position. The y-axis represents the log10 base transformed p-values. Single nucleotide*
388 *polymorphisms are represented by a single point. The red horizontal line indicates the genome*
389 *wide significance (GWS) threshold and the blue horizontal line indicates the suggestive wide*
390 *significance (SWS) threshold. Only genes above the GWS threshold are labeled, with the*
391 *exception of IGF1R. Marked genes are not drawn to scale.*

392 For breed average weight, 66 SNPs were observed above the SWS threshold and 45
393 SNPs above the GWS threshold, when including all individuals in our sample (Figure 7A).
394 Nearly half of the SNPs above the SWS threshold (30 SNPs) were found within genes on
395 chromosome 24. A single SNP, on chromosome 32 (position 11413601) had the strongest
396 association with weight (p $=3.24 \times 10^{-7}$). This was observed within *PKD2*, a gene that encodes
397 for protein kinase D2. Across all samples we identified several other SNPS to be above the SWS
398 threshold for weight: on chromosome 3, *CERS3* (7 SNPs); on chromosome 20, *ZXDC* (3 SNPs)
399 and *TXNRD3* (1 SNP); on chromosome 24, *SNTA1* (5 SNPs), *CBFA2T2* (19 SNPs), *ZNF341* (2
400 SNPs), and *CHMP4B* (4 SNPs); on chromosome 32, *SEPT11* (1 SNP), *CCNI* (1 SNP), *PKD2* (11
401 SNPs), *SGMS2* (4 SNPs) and *CYP2U1* (3 SNPs); on chromosome 35, *SPIDR* (5 SNPs) (Table
402 S14). The genes on chromosome 3, 20, 24, and 35 all had SNPs above the GWS threshold. The
403 only gene from chromosome 32 with SNPs above the GWS threshold was *PKD2*.

404 For small individuals, 31 SNPs were observed above the SWS threshold and 19 SNPS
405 above the GWS threshold (Figure 7B). Nearly a third of those above the SWS threshold (i.e. 20
406 SNPs) belong to chromosome 24. A single SNP on chromosome 24 (position 22660667) had the
407 strongest association with weight ($p = 5.88 \times 10^{-7}$) and is linked to the protein-coding gene
408 *SNTA1* (Figure 7B). For our subgroup of small individuals, we found SNPs above the SWS
409 threshold for weight to be linked to the following genes: on chromosome 3, *CERS3* (1 SNP); on
410 chromosome 9, *RPH3AL* (4 SNPs), PITPNA (4 SNPs), and SLC43A2 (2 SNPs); on chromosome
411 24, *ADAM33* (2 SNPs), *SIGLEC1* (2 SNPs), *HSPA12B* (1 SNP), *SNTA1* (4 SNPs), *CBFA2T2* (9
412 SNPs), *ZNF341* (1 SNP) and *CHMP4B* (1 SNP) (Table S14). SNPs on genes *PITPNA*,
413 *SLC43A2*, *SNTA1*, *CBFA2T2*, and *ZNF341* were above the GWS threshold. In contrast, for the
414 subgroup of large individuals, 10 SNPs were above the SWS threshold and 5 SNPs were above

415 the GWS threshold (Figure 7C). Only two genes reached significance in large individuals: on
416 chromosome 32, *PKD2* (4 SNPs) and on chromosome 35, *SPIDR* (6 SNPs). All the SNPs linked
417 to these two genes were above the SWS threshold for weight (Table S14). SNPs on both of these
418 genes were also above the GWS threshold. For an extended table containing all of the observed
419 genes associated with weight refer to Table S14.

420

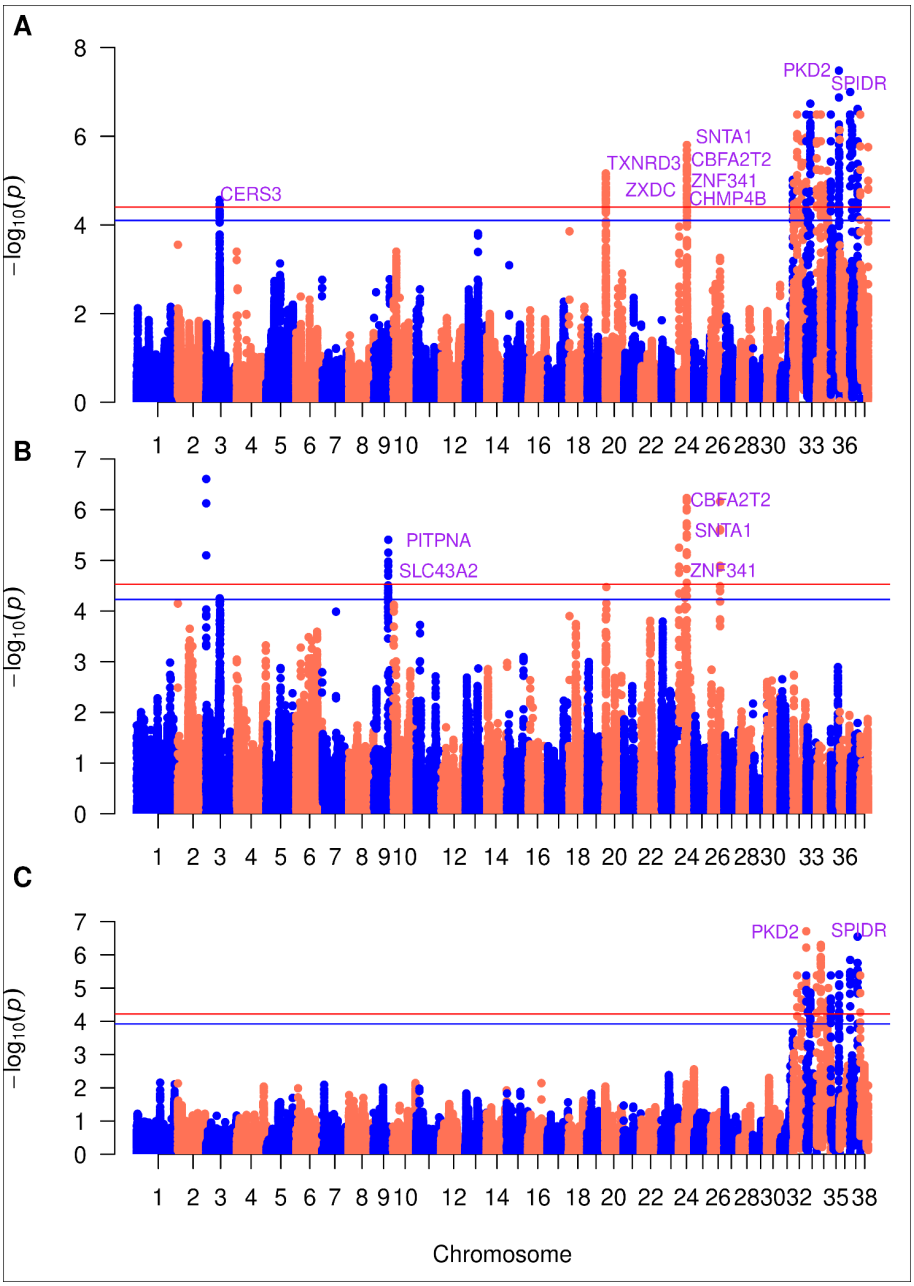


FIGURE 7. *Manhattan plots for GLMM-based GWAS using presence/absence of ROH for body weight. A. All individuals B. Small Individuals C. Large Individuals. The x-axis represents the genomic position. The y-axis represents the log10 base transformed p-values. Single nucleotide polymorphisms are represented by a single point. The red horizontal line indicates the genome wide significance (GWS) threshold and the blue horizontal line indicates the suggestive wide significance (SWS) threshold. Only genes above the GWS threshold are labeled. Marked genes are not drawn to scale.*

430 For breed average lifespan, 10 SNPs were observed above the SWS threshold and only 4 SNPs
431 were above the GWS threshold when tested across all individuals (Figure 8A). The strongest
432 associated SNP ($p = 2.21 \times 10^{-5}$) was located on chromosome 20 (position 789204) within the
433 gene *TXNRD3*. Across all samples we identified several other SNPS to be above the SWS
434 threshold for lifespan: on chromosome 20, *ZXDC* (2 SNPs) and *TXNRD3* (1 SNP) ; on
435 chromosome 32, *SEPT11* (1 SNP); on chromosome 33, *DRD3* (5 SNPs) and *TIGIT* (1 SNP)
436 (Table S15). All, but *SEPT11*, had a SNP above the GWS threshold. For our subgroup of small
437 individuals we only found a gene on chromosome 12 *RUNX2* (4 SNPs), to be above the SWS
438 threshold. (Figure 8B and Table S15). For our subgroup of large individuals, we found 9 SNPs to
439 be above the SWS threshold, all of which were also above the GWG threshold (Figure 8C). The
440 strongest associated SNP ($p = 1.17 \times 10^{-8}$) was located on chromosome 20 (position 789204)
441 within the gene *TXNRD3.* All genes with SNPs above the significance thresholds (*TXNRD3* (1
442 SNP), *ZXDC* (3 SNPs), *CFAP100* (4 SNPs), and *SLC41A3* (1 SNP)) were located on
443 chromosome 20 (Table S15). For an extended table containing all of the observed genes
444 associated with lifespan refer to Table S14.

445

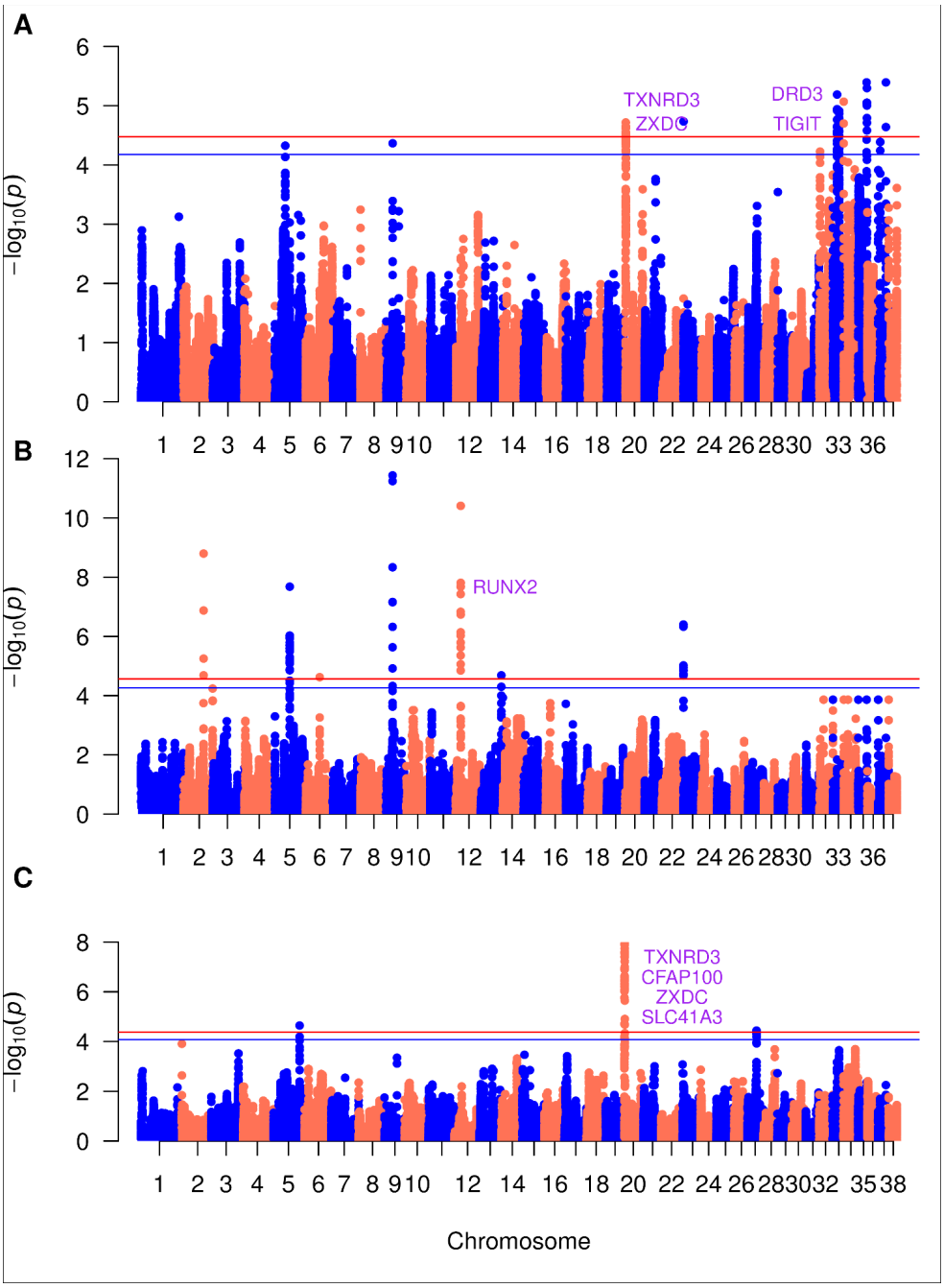


FIGURE 8. *Manhattan plots for GLMM-based GWAS using presence/absence of ROH for lifespan. A. All individuals B. Small Individuals C. Large Individuals. The x-axis represents the genomic position. The y-axis represents the log10 base transformed p-values. Single nucleotide polymorphisms are represented by a single point. The red horizontal line indicates the genome wide significance (GWS) threshold and the blue horizontal line indicates the suggestive wide significance (SWS) threshold. Only genes above the GWS threshold are labeled. Marked genes are not drawn to scale.*

454 # Discussion

455 Our work provides an in-depth analysis of patterns of runs of homozygosity (ROH) in
456 domesticated and non-domesticated dogs, highlighting how domestication and breed formation
457 has shaped genetic diversity in these animals. We also explore how ROH, quantified as the total
458 fraction of the genome within a run of homozygosity, F_{ROH}, can be used to uncover traits that are
459 not fully additive. We focus on ROH because these genomic segments reflect recent demography,
460 shared common ancestry due to domestication and breed formation, and can capture the effects of
461 recessive variation on complex traits (Mooney et al., 2021; Clark et al., 2019; Malawsky et al.,
462 2023; Purfield et al., 2017; Thompson, 2013).

463 For breed dogs, there is a large fraction of the genome within long ROH. Long ROH are a
464 result of very recent parental relatedness and are inextricably linked to this species' complex
465 evolutionary and domestication history. Previous research has shown that dogs originated from an
466 isolated wolf population(s), followed by recent strong artificial selection that drove breed
467 emergence (Boyko, 2011; Freedman & Wayne, 2017). The recent strong artificial selection during
468 breed formation resulted in low genetic diversity and high homozygosity within breeds, alongside
469 high phenotypic variance between breeds (Boyko et al., 2010). Artificial selection, in the case of
470 dogs, was achieved through inbreeding, and also led to the significant sharing of common genetic
471 variants (Gregory, 2009). Artificial selection for specific phenotypic characteristics during breed
472 establishment and subsequent breed standardization have resulted in homogeneity of both the
473 phenotypic and genotypic variation within a breed. Further, when looking between breeds,
474 previous work has shown that occasionally a dog with a desirable trait in one breed is also used to
475 introduce the same phenotype in another breed, which creates a network of genetic relatedness

476 through shared common haplotypes between breeds (Parker et al., 2017). In sum, because artificial
477 selection was often achieved through inbreeding during the domestication process, and breed
478 standards have reduced the effective population size of each breed, these two phenomena are both
479 reflected through shared haplotypic information between breeds and distinct patterns of
480 homozygosity within breeds.

481 In this work, we showed that domesticated breed dogs have unique patterns of ROH in their
482 genome, and on average carry more of their genome within long ROH than non-domesticated
483 village and Chinese indigenous dogs (Table 2). Conversely, the non-domesticated village dogs
484 and Chinese indigenous dogs have the lowest values for mean ROH distribution. Higher n_{ROH} is
485 indicative of a small population size and higher numbers of shorter n_{ROH} (i.e. class A) coincide
486 with bottlenecks (Ceballos et al., 2018). Taken together our results correspond with the higher
487 degree of relatedness among all domestic breed group individuals, and is expected given the
488 known high levels of inbreeding during breed establishment. Our results also highlight Chinese
489 indigenous dogs and village dogs as more outbred populations relative to breed dogs.
490 Specifically, we observed lower levels of mean n_{ROH} and mean s_{ROH} in Chinese indigenous dogs.
491 Lower n_{ROH} and mean s_{ROH} are indicative of a larger population size (Ceballos et al., 2018),
492 which is in line with the origins of the Chinese indigenous dogs. Previous work also highlighted
493 relaxed trait selection during establishment of the Chinese indigenous dog population (Q. Yang
494 et al., 2019), potentially this relaxed selection has resulted in overall higher levels of genetic
495 diversity. Moreover, the domesticated village dogs, which represented 13 different regions
496 around the world, had lower values for mean n_{ROH} and mean s_{ROH} (Table 1 and Table 2) relative
497 to breed dogs. This was expected since village dogs did not experience the additional bottleneck
498 during breed formation.

499 Turning to hotspots, we observed that there was some relationship with recombination rate per
500 chromosome and the density of ROH hotspots, however recombination does not fully explain all
501 of the patterns that we found (Figure S2, S3). For example, on many chromosomes, we observed
502 at least one ROH hotspot (across breed dogs and village dogs) in a region where recombination
503 rates peaks. This is also consistent with the conclusions from previous studies where researchers
504 suggest that recombination rates are not the only factor leading to ROH hotspots or cold spots
505 (Mooney et al., 2021; Pemberton & Szpiech, 2018). Our results suggest that the ROH patterns
506 we observe are instead largely driven by demography (domestication), artificial selection (breed
507 formation), and inbreeding. These more recent processes have resulted in unique ROH patterns
508 within breeds, sharing across breeds, and very little ROH sharing between breed dogs and village
509 dogs.

510 We also examined the relationship between ROH and complex traits, to identify traits that may
511 not be entirely additive. We observed that the relationship between breed average height and
512 F_{ROH} is positive when using all breed dogs (Figure 4). This positive effect size opposes recently
513 published work in human populations (Joshi et al., 2015; McQuillan et al., 2012; Swinford et al.,
514 2023), and may be a result of combining the effect sizes from multiple sizes of dogs (i.e. large
515 and small dogs). When we partition by weight, our results correspond to previous studies on
516 height in breed dogs (Hoopes et al., 2012). As expected, we found a positive β when testing for
517 the relationship between F_{ROH} and weight. Lastly, we observed a negative β with respect to the
518 relationship between F_{ROH} and lifespan, indicating an association with inbreeding due to
519 demography and survival to old age. This result is consistent with the findings in dogs where an
520 association with disease phenotypes was observed (Mooney et al., 2021). Further, we also
521 partitioned dogs by weight (Urfer et al., 2019) and observed a trend where large dog breeds have

522 a shorter lifespan ($\beta = -0.15$) and small dogs have a longer lifespan ($\beta = 0.11$), though our
523 results were not significant (Table S8). Previous studies have linked size with longevity; this
524 remains an interesting area for future work (Galis et al., 2007). Indeed, when we included
525 lifespan as a covariate when associating F_{ROH} and breed average height in all individuals we still
526 observed a nominally significant, but larger p-value, of 0.034. Further, the effect size of F_{ROH} on
527 height decreased from $\beta \sim 0.31$ to $\beta \sim 0.19$. We also observed a negative ($\beta \sim$ -0.83) and
528 significant relationship between height and lifespan ($p = 6.41 \times 10^{-41}$) that we expected to see.

529 Our GWAS using ROH revealed several significant associations between SNPs and the breed
530 average quantitative traits: height, weight, and lifespan (Table S13, Table S14, and Table S15).
531 When utilizing data for all individuals, we found three hits within *CERS3* (chromosome 3),
532 *CBFA2T2* (chromosome 24) and *SNTA1* (chromosome 24) to be above the Genome Wide
533 Significance (GWS) threshold for height. In previous studies, these three genes were found to
534 harbor genetic variants associated with human height (Kichaev et al., 2019; Yengo et al., 2022).
535 Additionally, human genetics studies have found *CERS3* to be linked with Body-Mass Index
536 (BMI) (Costa-Urrutia et al., 2019). We observed that many of the significant genic hits within
537 our subgroup of small individuals (*PITPNA*, *LRRC28*, *TTC23*, *HGF*, *SLC43A2*, and *VPS53*)
538 were also previously identified to be associated with human height (Guo et al., 2018; Yengo et
539 al., 2022). We also observed a relationship between *RPH3AL* and height. Though there is
540 limited research on *RPH3AL,* its role in exocytosis and the secretion of growth hormones in
541 humans has been noted (*RPH3AL Gene Information - The Human Protein Atlas*, n.d.).
542 Importantly, the signal within *IGF1R* replicated, and was above the SWS threshold. A mutation
543 in *IGF1R* was linked with height in small dog breeds, like Chihuahuas (Hoopes et al., 2012).
544 Lastly for small individuals, *COMMD6*, which was also identified in our GWAS, has primarily

been studied in relation to immune function, but, to our knowledge, no studies to date have reported any associations between this gene and height (*COMMD6 Gene Information - The Human Protein Atlas*, n.d.). In the subgroup of large individuals, the only gene above the SWS threshold was *PRR16* (chromosome 11); it has been previously associated with human height (Yengo et al., 2022).

Across all of our samples, our GWAS results pinpoint *PKD2* (chromosome 32) to be the most strongly associated with weight. Mouse models have shown the presence of *PKD2* has been linked to an increase in fat absorption, potentially promoting obesity (Trujillo-Viera et al., 2021). Several other genes, with significant associations in our study, have been linked to weight in other species. *SNTA1* was found to be closely associated with the Live Body Weight (LBW) of goats, which is a factor often used to assess livestock health (Selionova et al., 2022). Another gene, *CBFA2T2*, was found to regulate adipogenesis in both humans and mice, which can affect obesity when its expression is altered (Ali et al., 2013; Luo et al., 2020; Matulewicz et al., 2017). *TXNRD3* influences adipocyte differentiation through its involvement in the Wnt signaling pathway, a key process in the regulation of body fat and energy storage (S. Kang et al., 2007; Kipp et al., 2012; Martínez-Montes et al., 2017). *CERS3*, whose association was also noted in dog height, has been linked with Body Mass Index (BMI) in humans (Costa-Urrutia et al., 2019). Additionally, the genes *SPIDR*, and *ZNF341* have been linked to total body fat and bone mineral content, respectively (Blake et al., 2021; Shapses & Sukumar, 2012; Tachmazidou et al., 2017). While *CHMP4B* has been linked to the human birth weight, no studies have directly linked it to obesity and/or body mass. (*CHMP4B Gene Information - The Human Protein Atlas*, n.d.; Warrington et al., 2019). Similarly, *ZXDC* has primarily been investigated in the context of immune system regulation and cancer biology. To our knowledge, this is the first association

568 with weight thus far, which provides an additional avenue for future researchers to explore.
569 In our subgroup of small individuals, we identified similar associations between weight and the
570 genes *SNTA1*, *ZNF341*, and *CBFA2T2*. Additionally, we identified new associations specific to
571 this subgroup with *PITPNA* (also associated with height) and *SLC43A2*. Both of these genes
572 were shown to decrease body weight in knockout mice (Alb et al., 2003; Guetg et al., 2015). In
573 the subgroup of larger individuals we found two genes to be associated with weight, *PKD2* and
574 on *SPIDR*, both of which are associated with body fat.

575 For lifespan, we highlight several other associations of interest. When testing all samples, we
576 find significant SNPs linked to the genes: *TXNRD3*, *DRD3*, *TIGIT*, *SEPT11* and *ZXDC.*
577 *TXNRD3* has been previously associated with lower survival rates for various types of cancers
578 (Wu et al., 2021). However, to our knowledge, no previous studies have found an explicit
579 connection between *TXNRD3* and lifespan. The gene that encodes the D3 dopamine receptor,
580 *DRD3,* is associated with schizophrenia, which has too been linked with decreased life
581 expectancy (Hjorthøj et al., 2017; Y. Kang et al., 2023). We observed significant signals at
582 *TIGIT* and *SEPT11,* which both have been tied to the promotion of tumor growth and reduced
583 longevity (Chauvin et al., 2015; Fu et al., 2023; Johnston et al., 2014). *ZXDC* was previously
584 associated with cervical cancer metastasis, which has a significant negative impact on patient
585 survival (Mao et al., 2023). In small dogs, we saw an association with *RUNX2*, which has also
586 been associated with multiple cancers and poor patient prognosis (Zhao et al., 2021). In large
587 individuals, our GWAS revealed *SLC41A3* to be associated with lifespan. A relationship between
588 this gene and liver hepatocellular carcinoma, which is the second primary contributor of
589 carcinoma-associated death, has been established (Q. Chang et al., 2021). Another gene found in

590 the group of large individuals is *CFAP100*, which to our knowledge, has not previously been
591 linked to lifespan.

592 In humans, there have been 25,551 associations with height, 2,233 associations with weight, and
593 664 associations with lifespan as of 2023 (Sollis et al., 2023). After correcting for inflation, we
594 recover a total of 50 associations with height, 84 associations with weight, and 21 associations
595 with lifespan. This corresponds to previous work, which has suggested that dogs have a more
596 simplified complex trait architecture than humans. (Boyko et al., 2010; Freedman et al., 2016).
597 Importantly, we conducted an additional test to validate our ROH based GWAS by attempting to
598 replicate a previously identified peak that was known to be homozygous on chromosome 13
599 within RSPO2 for furnishings in dogs (Plassais et al., 2019). We were able to identify the same
600 peak on chromosome 13 for RSPO2 (Figures S27 and S28), though it did not meet genome-wide
601 or suggestive significance after p-values were corrected. In addition to the RSPO2 peak, we
602 observed additional peaks that had not been previously identified. Replicating this peak suggests
603 our approach which aimed at capturing slightly different information is powerful and worth-
604 while. Notably, similar power in human GWAS requires hundreds of thousands or even a million
605 individuals (Yengo et al., 2022). Here, we captured slightly different information (non-
606 additivity) when using ROH-mapping GWAS, and we were able to capture this information with
607 less than 1000 dogs. This is due to the unique domestication and breed formation bottlenecks that
608 dogs experienced that simplified complex trait architecture (Boyko et al., 2010; Hayward et al.,
609 2016; Plassais et al., 2019). Overall, our GWAS and association tests suggest that quantitative
610 traits in dogs are polygenic, and some traits may not be entirely additive. The newly identified
611 genes, which have not been observed in previous studies, could be used as candidates in future
612 functional exploration.

613 There are some limitations in the work that we would like to highlight. First, the phenotypic
614 values used in the analysis were calculated based on the breed average, restricting us to study
615 associations between breeds. In other words, the within breed values would not yield any
616 significant results as the expected association between the trait and F_{ROH} would be 0, due to
617 using a breed average phenotype. Using the breed average has been previously validated (Boyko
618 et al., 2010; Hayward et al., 2016; Jones et al., 2008; Plassais et al., 2017, 2019; Rimbault et al.,
619 2013) but the assumption does come with caveats. For example, when using a breed average
620 value, we are ignoring variance among individuals within a breed. This variance within the breed
621 is quite small due to the strict standards of the AKC. To this end, we mathematically show
622 (Supplementary Text) that as long as the within breed variance is small the effect size computed
623 from individual level data will be equivalent to the effect size from using the average. However,
624 bias can be introduced in some cases when aggregating across breeds (Figure S29). Secondly,
625 breed group categorization was based on neighbor joining trees and AKC, which could introduce
626 errors when grouping some breeds. Third, observed that there was a concentration of peaks in the
627 last few chromosomes. After eliminating SNP density issues (Figure S30, Table S16), we believe
628 that the large number of peaks on the smaller chromosomes might be due to stronger linkage
629 disequilibrium in the smaller chromosomes amplifying the strength of the signal. Alternatively,
630 some chromosomes such as chromosome 32 have a higher GC content when examining the
631 CanFam3.1 assembly (C. Wang et al., 2021), which could introduce noise into our results.

632 In sum, our work has further elucidated the genetic structure of quantitative trait architecture in
633 dogs. We identified significant relationships between F_{ROH} and multiple phenotypic traits,
634 suggesting that domestication and artificial selection played a role in shaping complex non-disease

traits. Our ROH-based association tests and GWAS provide new avenues to explore non-additivity in domesticated species while further elucidating complex trait architecture.

Author Contributions

ZAS and JAM designed and supervised the research project. S performed the research and analyzed data. SN reviewed domestication literature for writing and editing of the manuscript. SH generated the proof and plots for the effect of using breed averages. S along with the assistance of SN, ZAS and JAM wrote the paper.

Acknowledgements

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM146926 (ZAS and S). This work was also supported by Eberly College of Science Startup Fund (ZAS and S). Computations for this research were performed using the Pennsylvania State University's Institute for Computational Data Sciences' Roar supercomputer. S would also like to thank the entire Szpiech Lab for their invaluable support. JAM was supported by the startup funds from Dornsife College of Letters, Arts and Sciences through the Department of Quantitative and Computational Biology and the USC WiSE Gabilan Assistant Professorship.

Data Accessibility Statement

Whole Genome Sequencing is available on NCBI (Accession number: PRJNA448733). The source phenotype data was obtained from (Plassais et al., 2019). All other data are contained within the article and its supplementary information.

Bibliography

Alb, J. G., Cortese, J. D., Phillips, S. E., Albin, R. L., Nagy, T. R., Hamilton, B. A., & Bankaitis, V. A. (2003). Mice lacking phosphatidylinositol transfer protein-α exhibit spinocerebellar degeneration, intestinal and hepatic steatosis, and hypoglycemia. *Journal of Biological Chemistry*, *278*(35), 33501–33518. https://doi.org/10.1074/jbc.M303591200

Ali, A. T., Hochfeld, W. E., Myburgh, R., & Pepper, M. S. (2013). Adipocyte and adipogenesis. *European Journal of Cell Biology*, *92*(6–7), 229–236. https://doi.org/10.1016/j.ejcb.2013.06.001

Bacolod, M. D., Schemmann, G. S., Wang, S., Shattock, R., Giardina, S. F., Zeng, Z., Shia, J., Stengel, R. F., Gerry, N., Hoh, J., Kirchhoff, T., Gold, B., Christman, M. F., Offit, K., Gerald, W. L., Notterman, D. A., Ott, J., Paty, P. B., & Barany, F. (2008). The signatures of autozygosity among patients with colorectal cancer. *Cancer Research*, *68*(8), 2610–2621. https://doi.org/10.1158/0008-5472.CAN-07-5250

Bannasch, D., Famula, T., Donner, J., Anderson, H., Honkanen, L., Batcher, K., Safra, N., Thomasy, S., & Rebhun, R. (2021). The effect of inbreeding, body size and morphology on health in dog breeds. *Canine Medicine and Genetics*, *8*(1), 12. https://doi.org/10.1186/s40575-021-00111-4

Barton, A. R., Hujoel, M. L. A., Mukamel, R. E., Sherman, M. A., & Loh, P.-R. (2022). A spectrum of recessiveness among Mendelian disease variants in UK Biobank. *The American Journal of Human Genetics*, *109*(7), 1298–1307. https://doi.org/10.1016/j.ajhg.2022.05.008

Benecke, N. (1987). Studies on early dog remains from Northern Europe. *Journal of Archaeological Science*, *14*(1), 31–49. https://doi.org/10.1016/S0305-4403(87)80004-3

Bergström, A., Stanton, D. W. G., Taron, U. H., Frantz, L., Sinding, M.-H. S., Ersmark, E., Pfrengle, S., Cassatt-Johnstone, M., Lebrasseur, O., Girdland-Flink, L., Fernandes, D.

M., Ollivier, M., Speidel, L., Gopalakrishnan, S., Westbury, M. V., Ramos-Madrigal, J., Feuerborn, T. R., Reiter, E., Gretzinger, J., … Skoglund, P. (2022). Grey wolf genomic history reveals a dual ancestry of dogs. *Nature*, *607*(7918), 313–320. https://doi.org/10.1038/s41586-022-04824-9

Blake, J. A., Baldarelli, R., Kadin, J. A., Richardson, J. E., Smith, C. L., Bult, C. J., the Mouse Genome Database Group, Anagnostopoulos, A. V., Beal, J. S., Bello, S. M., Blodgett, O., Butler, N. E., Campbell, J., Christie, K. R., Corbani, L. E., Dolan, M. E., Drabkin, H. J., Flores, M., Giannatto, S. L., … Zhu, Y. 'Sophia.' (2021). Mouse Genome Database (MGD): Knowledgebase for mouse–human comparative biology. *Nucleic Acids Research*, *49*(D1), D981–D987. https://doi.org/10.1093/nar/gkaa1083

Boyko, A. R. (2011). The domestic dog: Man's best friend in the genomic era. *Genome Biology*, *12*(2), 216. https://doi.org/10.1186/gb-2011-12-2-216

Boyko, A. R., Quignon, P., Li, L., Schoenebeck, J. J., Degenhardt, J. D., Lohmueller, K. E., Zhao, K., Brisbin, A., Parker, H. G., vonHoldt, B. M., Cargill, M., Auton, A., Reynolds, A., Elkahloun, A. G., Castelhano, M., Mosher, D. S., Sutter, N. B., Johnson, G. S., Novembre, J., … Ostrander, E. A. (2010). A simple genetic architecture underlies morphological variation in dogs. *PLOS Biology*, *8*(8), e1000451. https://doi.org/10.1371/journal.pbio.1000451

Cadieu, E., Neff, M. W., Quignon, P., Walsh, K., Chase, K., Parker, H. G., VonHoldt, B. M., Rhue, A., Boyko, A., Byers, A., Wong, A., Mosher, D. S., Elkahloun, A. G., Spady, T. C., André, C., Lark, K. G., Cargill, M., Bustamante, C. D., Wayne, R. K., & Ostrander, E. A. (2009). Coat variation in the domestic dog is governed by variants in three genes. *Science*, *326*(5949), 150–153. https://doi.org/10.1126/science.1177808

Candille, S. I., Kaelin, C. B., Cattanach, B. M., Yu, B., Thompson, D. A., Nix, M. A., Kerns, J. A., Schmutz, S. M., Millhauser, G. L., & Barsh, G. S. (2007). A β-Defensin mutation causes black coat color in domestic dogs. *Science*, *318*(5855), 1418–1423.

https://doi.org/10.1126/science.1147880

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, *19*(4), 220–234. https://doi.org/10.1038/nrg.2017.109

Cecchi, F., Paci, G., Spaterna, A., & Ciampolini, R. (2013). Genetic variability in *Bracco Italiano* dog breed assessed by pedigree data. *Italian Journal of Animal Science*, *12*(3), e54. https://doi.org/10.4081/ijas.2013.e54

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Chang, Q., Xu, Y., Wang, J., Jing, H., Rao, L., Tang, W., Zhang, Z., & Wu, X. (2021). SLC41A3 exhibits as a carcinoma biomarker and promoter in liver hepatocellular carcinoma. *Computational and Mathematical Methods in Medicine*, *2021*, 1–9. https://doi.org/10.1155/2021/8556888

Chauvin, J.-M., Pagliano, O., Fourcade, J., Sun, Z., Wang, H., Sander, C., Kirkwood, J. M., Chen, T. T., Maurer, M., Korman, A. J., & Zarour, H. M. (2015). TIGIT and PD-1 impair tumor antigen–specific CD8+ T cells in melanoma patients. *Journal of Clinical Investigation*, *125*(5), 2046–2058. https://doi.org/10.1172/JCI80445

Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., Szpiro, A. A., Chen, W., Brehm, J. M., Celedón, J. C., Redline, S., Papanicolaou, G. J., Thornton, T. A., Laurie, C. C., Rice, K., & Lin, X. (2016). Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *The American Journal of Human Genetics*, *98*(4), 653–666. https://doi.org/10.1016/j.ajhg.2016.02.012

CHMP4B gene information—The Human Protein Atlas. (n.d.). Retrieved October 30, 2024, from https://www.proteinatlas.org/ENSG00000101421-CHMP4B/summary/gene

Chundru, V. K., Zhang, Z., Walter, K., Lindsay, S. J., Danecek, P., Eberhardt, R. Y., Gardner, E.

J., Malawsky, D. S., Wigdor, E. M., Torene, R., Retterer, K., Wright, C. F., Ólafsdóttir, H., Guillen Sacoto, M. J., Ayaz, A., Akbeyaz, I. H., Türkdoğan, D., Al Balushi, A. I., Bertoli-Avella, A., … Martin, H. C. (2024). Federated analysis of autosomal recessive coding variants in 29,745 developmental disorder patients from diverse populations. *Nature Genetics*, *56*(10), 2046–2053. https://doi.org/10.1038/s41588-024-01910-8

Clark, D. W., Okada, Y., Moore, K. H. S., Mason, D., Pirastu, N., Gandin, I., Mattsson, H., Barnes, C. L. K., Lin, K., Zhao, J. H., Deelen, P., Rohde, R., Schurmann, C., Guo, X., Giulianini, F., Zhang, W., Medina-Gomez, C., Karlsson, R., Bao, Y., … Wilson, J. F. (2019). Associations of autozygosity with a broad range of human phenotypes. *Nature Communications*, *10*(1), 4957. https://doi.org/10.1038/s41467-019-12283-6

COMMD6 gene information—The Human Protein Atlas. (n.d.). Retrieved October 30, 2024, from https://www.proteinatlas.org/ENSG00000188243-COMMD6/summary/gene

Costa-Urrutia, P., Colistro, V., Jiménez-Osorio, A. S., Cárdenas-Hernández, H., Solares-Tlapechco, J., Ramirez-Alcántara, M., Granados, J., Ascencio-Montiel, I. D. J., & Rodríguez-Arellano, M. E. (2019). Genome-wide association study of body mass index and body fat in Mexican-Mestizo children. *Genes*, *10*(11), 945. https://doi.org/10.3390/genes10110945

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Darwin, C. (1868). *The variation of animals and plants under domestication: Vol. I*. John Murray.

Davis, S. J. M., & Valla, F. R. (1978). Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. *Nature*, *276*(5688), 608–610. https://doi.org/10.1038/276608a0

Dog Breeds—Types Of Dogs. (n.d.). American Kennel Club. Retrieved October 14, 2024, from https://www.akc.org/dog-breeds/

Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A. S., Schweizer, R. M., Ramirez, O., Pollinger,

759 J., Galaverni, M., Ortega Del-Vecchyo, D., Du, L., Zhang, W., Zhang, Z., Xing, J., Vilà,
760 C., Marques-Bonet, T., Godinho, R., Yue, B., & Wayne, R. K. (2016). Worldwide patterns
761 of genomic variation and admixture in gray wolves. *Genome Research*, *26*(2), 163–173.
762 https://doi.org/10.1101/gr.197517.115
763 Fareed, M., & Afzal, M. (2014). Evidence of inbreeding depression on height, weight, and body
764 mass index: A population-based child cohort study. *American Journal of Human Biology*,
765 *26*(6), 784–795. https://doi.org/10.1002/ajhb.22599
766 Freedman, A. H., Gronau, I., Schweizer, R. M., Vecchyo, D. O.-D., Han, E., Silva, P. M.,
767 Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O.,
768 Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffen, E., Kusak, J., Boyko, A. R., …
769 Novembre, J. (2014). Genome sequencing highlights the dynamic early history of dogs.
770 *PLOS Genetics*, *10*(1), e1004016. https://doi.org/10.1371/journal.pgen.1004016
771 Freedman, A. H., Lohmueller, K. E., & Wayne, R. K. (2016). Evolutionary history, selective
772 sweeps, and deleterious variation in the dog. *Annual Review of Ecology, Evolution, and*
773 *Systematics*, *47*(1), 73–96. https://doi.org/10.1146/annurev-ecolsys-121415-032155
774 Freedman, A. H., & Wayne, R. K. (2017). Deciphering the origin of dogs: From fossils to
775 genomes. *Annual Review of Animal Biosciences*, *5*(1), 281–307.
776 https://doi.org/10.1146/annurev-animal-022114-110937
777 Fridman, H., Khazeeva, G., Levy-Lahad, E., Gilissen, C., & Brunner, H. G. (2024). *Reproductive*
778 *and cognitive effects in carriers of recessive pathogenic variants*.
779 https://doi.org/10.1101/2024.09.30.615774
780 Fu, L., Wang, X., Yang, Y., Chen, M., Kuerban, A., Liu, H., Dong, Y., Cai, Q., Ma, M., & Wu, X.
781 (2023). Septin11 promotes hepatocellular carcinoma cell motility by activating RhoA to
782 regulate cytoskeleton and cell adhesion. *Cell Death & Disease*, *14*(4), 280.
783 https://doi.org/10.1038/s41419-023-05726-y
784 Galis, F., Van Der Sluijs, I., Van Dooren, T. J. M., Metz, J. A. J., & Nussbaumer, M. (2007). Do

large dogs die young? *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *308B*(2), 119–126. https://doi.org/10.1002/jez.b.21116

Gray, A. P. (1972). *Mammalian hybrids: A check-list with bibliography* (2d [rev.] ed). Commonwealth Agricultural Bureaux.

Gregory, T. R. (2009). Artificial selection and domestication: Modern lessons from Darwin's enduring analogy. *Evolution: Education and Outreach*, *2*(1), 5–27. https://doi.org/10.1007/s12052-008-0114-z

Guagnin, M., Perri, A. R., & Petraglia, M. D. (2018). Pre-Neolithic evidence for dog-assisted hunting strategies in Arabia. *Journal of Anthropological Archaeology*, *49*, 225–236. https://doi.org/10.1016/j.jaa.2017.10.003

Guetg, A., Mariotta, L., Bock, L., Herzog, B., Fingerhut, R., Camargo, S. M. R., & Verrey, F. (2015). Essential amino acid transporter Lat4 (*Slc43a2*) is required for mouse development. *The Journal of Physiology*, *593*(5), 1273–1289. https://doi.org/10.1113/jphysiol.2014.283960

Guo, M. H., Hirschhorn, J. N., & Dauber, A. (2018). Insights and implications of genome-wide association studies of height. *The Journal of Clinical Endocrinology & Metabolism*, *103*(9), 3155–3168. https://doi.org/10.1210/jc.2018-01126

Hayward, J. J., Castelhano, M. G., Oliveira, K. C., Corey, E., Balkman, C., Baxter, T. L., Casal, M. L., Center, S. A., Fang, M., Garrison, S. J., Kalla, S. E., Korniliev, P., Kotlikoff, M. I., Moise, N. S., Shannon, L. M., Simpson, K. W., Sutter, N. B., Todhunter, R. J., & Boyko, A. R. (2016). Complex disease and phenotype mapping in the domestic dog. *Nature Communications*, *7*(1), 10460. https://doi.org/10.1038/ncomms10460

Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: A systematic review and meta-analysis. *The Lancet Psychiatry*, *4*(4), 295–301. https://doi.org/10.1016/S2215-0366(17)30078-0

Hoopes, B. C., Rimbault, M., Liebers, D., Ostrander, E. A., & Sutter, N. B. (2012). The insulin-

like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs. *Mammalian Genome*, *23*(11–12), 780–790. https://doi.org/10.1007/s00335-012-9417-z

Johnson, E. C., Evans, L. M., & Keller, M. C. (2018). Relationships between estimated autozygosity and complex traits in the UK Biobank. *PLOS Genetics*, *14*(7), e1007556. https://doi.org/10.1371/journal.pgen.1007556

Johnston, R. J., Comps-Agrar, L., Hackney, J., Yu, X., Huseni, M., Yang, Y., Park, S., Javinal, V., Chiu, H., Irving, B., Eaton, D. L., & Grogan, J. L. (2014). The immunoreceptor TIGIT regulates antitumor and antiviral CD8 + T cell effector function. *Cancer Cell*, *26*(6), 923–937. https://doi.org/10.1016/j.ccell.2014.10.018

Jones, P., Chase, K., Martin, A., Davern, P., Ostrander, E. A., & Lark, K. G. (2008). Single-nucleotide-polymorphism-based association mapping of dog stereotypes. *Genetics*, *179*(2), 1033–1044. https://doi.org/10.1534/genetics.108.087866

Joshi, P. K., Esko, T., Mattsson, H., Eklund, N., Gandin, I., Nutile, T., Jackson, A. U., Schurmann, C., Smith, A. V., Zhang, W., Okada, Y., Stančáková, A., Faul, J. D., Zhao, W., Bartz, T. M., Concas, M. P., Franceschini, N., Enroth, S., Vitart, V., … Wilson, J. F. (2015). Directional dominance on stature and cognition in diverse human populations. *Nature*, *523*(7561), 459–462. https://doi.org/10.1038/nature14618

Kang, S., Bennett, C. N., Gerin, I., Rapp, L. A., Hankenson, K. D., & MacDougald, O. A. (2007). Wnt signaling stimulates osteoblastogenesis of mesenchymal precursors by suppressing CCAAT/Enhancer-binding protein α and peroxisome proliferator-activated receptor γ. *Journal of Biological Chemistry*, *282*(19), 14515–14524. https://doi.org/10.1074/jbc.M700030200

Kang, Y., Zhang, Y., Huang, K., & Wang, Z. (2023). The genetic influence of the DRD3 rs6280 polymorphism (Ser9Gly) on functional connectivity and gray matter volume of the hippocampus in patients with first-episode, drug-naïve schizophrenia. *Behavioural Brain Research*, *437*, 114124. https://doi.org/10.1016/j.bbr.2022.114124

Keller, M. C., Simonson, M. A., Ripke, S., Neale, B. M., Gejman, P. V., Howrigan, D. P., Lee, S. H., Lencz, T., Levinson, D. F., Sullivan, P. F., & The Schizophrenia Psychiatric Genome-Wide Association Study Consortium. (2012). Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. *PLoS Genetics*, *8*(4), e1002656. https://doi.org/10.1371/journal.pgen.1002656

Kichaev, G., Bhatia, G., Loh, P.-R., Gazal, S., Burch, K., Freund, M. K., Schoech, A., Pasaniuc, B., & Price, A. L. (2019). Leveraging polygenic functional enrichment to improve GWAS power. *The American Journal of Human Genetics*, *104*(1), 65–75. https://doi.org/10.1016/j.ajhg.2018.11.008

Kim, K., Jung, J., Caetano-Anollés, K., Sung, S., Yoo, D., Choi, B.-H., Kim, H.-C., Jeong, J.-Y., Cho, Y.-M., Park, E.-W., Choi, T.-J., Park, B., Lim, D., & Kim, H. (2018). Artificial selection increased body weight but induced increase of runs of homozygosity in Hanwoo cattle. *PLOS ONE*, *13*(3), e0193701. https://doi.org/10.1371/journal.pone.0193701

Kipp, A. P., Müller, M. F., Göken, E. M., Deubel, S., & Brigelius-Flohé, R. (2012). The selenoproteins GPx2, TrxR2 and TrxR3 are regulated by Wnt signalling in the intestinal epithelium. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1820*(10), 1588–1596. https://doi.org/10.1016/j.bbagen.2012.05.016

Larson, G., & Bradley, D. G. (2014). How How much is that in dog years? The advent of canine population genomics. *PLoS Genetics*, *10*(1), e1004093. https://doi.org/10.1371/journal.pgen.1004093

Li, M.-X., Yeung, J. M. Y., Cherny, S. S., & Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, *131*(5), 747–756. https://doi.org/10.1007/s00439-011-1118-2

Luo, J., Dou, L., Yang, Z., Zhou, Z., & Huang, H. (2020). CBFA2T2 promotes adipogenic

differentiation of mesenchymal stem cells by regulating CEBPA. *Biochemical and Biophysical Research Communications*, *529*(2), 133–139. https://doi.org/10.1016/j.bbrc.2020.05.120

Lynch, M. T., Maloney, K. A., Xu, H., Perry, J. A., Center, R. G., Shuldiner, A. R., & Mitchell, B. D. (2023). Associations of genome-wide and regional autozygosity with 96 complex traits in old order Amish. *BMC Genomics*, *24*(1), 134. https://doi.org/10.1186/s12864-023-09208-5

Malawsky, D. S., Van Walree, E., Jacobs, B. M., Heng, T. H., Huang, Q. Q., Sabir, A. H., Rahman, S., Sharif, S. M., Khan, A., Mirkov, M. U., Kuwahara, H., Gao, X., Alkuraya, F. S., Posthuma, D., Newman, W. G., Griffiths, C. J., Mathur, R., Van Heel, D. A., Finer, S., … Martin, H. C. (2023). Influence of autozygosity on common disease risk across the phenotypic spectrum. *Cell*, *186*(21), 4514-4527.e14. https://doi.org/10.1016/j.cell.2023.08.028

Mao, Y., Jiang, X., Guo, P., Ouyang, Y., Chen, X., Xia, M., Wu, L., Tang, Z., Liang, T., Li, Y., & He, M. (2023). ZXDC enhances cervical cancer metastasis through IGF2BP3-mediated activation of RhoA/ROCK signaling. *iScience*, *26*(8), 107447. https://doi.org/10.1016/j.isci.2023.107447

Martínez-Montes, A. M., Muiños-Bühl, A., Fernández, A., Folch, J. M., Ibáñez-Escriche, N., & Fernández, A. I. (2017). Deciphering the regulation of porcine genes influencing growth, fatness and yield-related traits through genetical genomics. *Mammalian Genome*, *28*(3–4), 130–142. https://doi.org/10.1007/s00335-016-9674-3

Mastrangelo, S., Biscarini, F., Auzino, B., Ragatzu, M., Spaterna, A., & Ciampolini, R. (2018). Genome-wide diversity and runs of homozygosity in the “Braque Français, type Pyrénées” dog breed. *BMC Research Notes*, *11*(1), 13. https://doi.org/10.1186/s13104-017-3112-9

Matulewicz, N., Stefanowicz, M., Nikołajuk, A., & Karczewska-Kupczewska, M. (2017). Markers

of adipogenesis, but not inflammation, in adipose tissue are independently related to insulin sensitivity. *The Journal of Clinical Endocrinology & Metabolism*, *102*(8), 3040–3049. https://doi.org/10.1210/jc.2017-00597

McQuillan, R. (2009). *Homozygosity, inbreeding and health in European populations*. https://era.ed.ac.uk/handle/1842/5946

McQuillan, R., Eklund, N., Pirastu, N., Kuningas, M., McEvoy, B. P., Esko, T., Corre, T., Davies, G., Kaakinen, M., Lyytikäinen, L.-P., Kristiansson, K., Havulinna, A. S., Gögele, M., Vitart, V., Tenesa, A., Aulchenko, Y., Hayward, C., Johansson, Å., Boban, M., … on behalf of the ROHgen Consortium. (2012). Evidence of inbreeding depression on human height. *PLoS Genetics*, *8*(7), e1002655. https://doi.org/10.1371/journal.pgen.1002655

Meadows, J. R. S., Kidd, J. M., Wang, G.-D., Parker, H. G., Schall, P. Z., Bianchi, M., Christmas, M. J., Bougiouri, K., Buckley, R. M., Hitte, C., Nguyen, A. K., Wang, C., Jagannathan, V., Niskanen, J. E., Frantz, L. A. F., Arumilli, M., Hundi, S., Lindblad-Toh, K., Ginja, C., … Ostrander, E. A. (2023). Genome sequencing of 2000 canids by the Dog10K consortium advances the understanding of demography, genome function and architecture. *Genome Biology*, *24*(1), 187. https://doi.org/10.1186/s13059-023-03023-7

Mooney, J. A., Yohannes, A., & Lohmueller, K. E. (2021). The impact of identity by descent on fitness and disease in dogs. *Proceedings of the National Academy of Sciences*, *118*(16), e2019116118. https://doi.org/10.1073/pnas.2019116118

Moreno-Grau, S., Fernández, M. V., De Rojas, I., Garcia-González, P., Hernández, I., Farias, F., Budde, J. P., Quintela, I., Madrid, L., González-Pérez, A., Montrreal, L., Alarcón-Martín, E., Alegret, M., Maroñas, O., Pineda, J. A., Macías, J., The GR@ACE study group, Abdelnour, C., Aguilera, N., … for the Alzheimer's Disease Neuroimaging Initiative. (2021). Long runs of homozygosity are associated with Alzheimer's disease. *Translational Psychiatry*, *11*(1), 142. https://doi.org/10.1038/s41398-020-01145-1

Morey, D. F. (2006). Burying key evidence: The social bond between dogs and people. *Journal*

of Archaeological Science, *33*(2), 158–175. https://doi.org/10.1016/j.jas.2005.07.009

Nagasawa, M., Mitsui, S., En, S., Ohtani, N., Ohta, M., Sakuma, Y., Onaka, T., Mogi, K., & Kikusui, T. (2015). Oxytocin-gaze positive loop and the coevolution of human-dog bonds. *Science*, *348*(6232), 333–336. https://doi.org/10.1126/science.1261022

Nobis, G. (1979). Der älteste Haushund lebte vor 14.000 Jahren. *Umschau in Wissenschaft Und Technik*, *79*(19), 610.

Ostrander, E. A., Wang, G.-D., Larson, G., vonHoldt, B. M., Davis, B. W., Jagannathan, V., Hitte, C., Wayne, R. K., Zhang, Y.-P., Dog10K Consortium, André, C., Axelsson, E., Boyko, A., Davis, B. W., Forman, O., Frantz, L., Hitte, C., Jagannathan, V., Karlsson, E., … Zhang, *Ya-Ping. (2019). Dog10K: An international sequencing effort to advance studies of canine domestication, phenotypes and health. *National Science Review*, *6*(4), 810–824. https://doi.org/10.1093/nsr/nwz049

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A. -b., Luo, L.-Y., Angleby, H., Ardalan, A., Ekstrom, C., Skollermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.-P., & Savolainen, P. (2009). mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Molecular Biology and Evolution*, *26*(12), 2849–2864. https://doi.org/10.1093/molbev/msp195

Parker, H. G., Dreger, D. L., Rimbault, M., Davis, B. W., Mullen, A. B., Carpintero-Ramirez, G., & Ostrander, E. A. (2017). Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Reports*, *19*(4), 697–708. https://doi.org/10.1016/j.celrep.2017.03.079

Parker, H. G., VonHoldt, B. M., Quignon, P., Margulies, E. H., Shao, S., Mosher, D. S., Spady, T. C., Elkahloun, A., Cargill, M., Jones, P. G., Maslen, C. L., Acland, G. M., Sutter, N. B., Kuroki, K., Bustamante, C. D., Wayne, R. K., & Ostrander, E. A. (2009). An Expressed *Fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science*, *325*(5943), 995–998. https://doi.org/10.1126/science.1173275

Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics*, *91*(2), 275–292. https://doi.org/10.1016/j.ajhg.2012.06.014

Pemberton, T. J., & Szpiech, Z. A. (2018). Relationship between deleterious variation, genomic autozygosity, and disease risk: Insights from the 1000 genomes project. *The American Journal of Human Genetics*, *102*(4), 658–675. https://doi.org/10.1016/j.ajhg.2018.02.013

Perri, A. R., Feuerborn, T. R., Frantz, L. A. F., Larson, G., Malhi, R. S., Meltzer, D. J., & Witt, K. E. (2021). Dog domestication and the dual dispersal of people and dogs into the Americas. *Proceedings of the National Academy of Sciences*, *118*(6), e2010083118. https://doi.org/10.1073/pnas.2010083118

Plassais, J., Kim, J., Davis, B. W., Karyadi, D. M., Hogan, A. N., Harris, A. C., Decker, B., Parker, H. G., & Ostrander, E. A. (2019). Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nature Communications*, *10*(1), 1489. https://doi.org/10.1038/s41467-019-09373-w

Plassais, J., Rimbault, M., Williams, F. J., Davis, B. W., Schoenebeck, J. J., & Ostrander, E. A. (2017). Analysis of large versus small dogs reveals three genes on the canine X chromosome associated with body weight, muscling and back fat thickness. *PLOS Genetics*, *13*(3), e1006661. https://doi.org/10.1371/journal.pgen.1006661

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11.

Purfield, D. C., McParland, S., Wall, E., & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLOS ONE*, *12*(5), e0176780. https://doi.org/10.1371/journal.pone.0176780

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rimbault, M., Beale, H. C., Schoenebeck, J. J., Hoopes, B. C., Allen, J. J., Kilroy-Glynn, P., Wayne, R. K., Sutter, N. B., & Ostrander, E. A. (2013). Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Research*, *23*(12), 1985–1995. https://doi.org/10.1101/gr.157339.113

RPH3AL gene information—The Human Protein Atlas. (n.d.). Retrieved October 30, 2024, from https://www.proteinatlas.org/ENSG00000181031-RPH3AL/summary/gene

Sams, A. J., & Boyko, A. R. (2019). Fine-scale resolution of runs of homozygosity reveal patterns of inbreeding and substantial overlap with recessive disease genotypes in domestic dogs. *G3 Genes|Genomes|Genetics*, *9*(1), 117–123. https://doi.org/10.1534/g3.118.200836

Savolainen, P., Zhang, Y., Luo, J., Lundeberg, J., & Leitner, T. (2002). Genetic evidence for an East Asian origin of domestic dogs. *Science*, *298*(5598), 1610–1613. https://doi.org/10.1126/science.1073906

Selionova, M., Aibazov, M., Mamontova, T., Malorodov, V., Sermyagin, A., Zinovyeva, N., & Easa, A. A. (2022). Genome-wide association study of live body weight and body conformation traits in young Karachai goats. *Small Ruminant Research*, *216*, 106836. https://doi.org/10.1016/j.smallrumres.2022.106836

Shapses, S. A., & Sukumar, D. (2012). Bone metabolism in obesity and weight loss. *Annual Review of Nutrition*, *32*(1), 287–309. https://doi.org/10.1146/annurev.nutr.012809.104655

Shriner, D., Adeyemo, A., & Rotimi, C. N. (2011). Joint Ancestry and Association Testing in Admixed Individuals. *PLoS Computational Biology*, *7*(12), e1002325. https://doi.org/10.1371/journal.pcbi.1002325

Sofer, T., Zheng, X., Laurie, C. A., Gogarten, S. M., Brody, J. A., Conomos, M. P., Bis, J. C., Thornton, T. A., Szpiro, A., O'Connell, J. R., Lange, E. M., Gao, Y., Cupples, L. A., Psaty, B. M., NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, Abe,

N., Abecasis, G., Aguet, F., Albert, C., … Rice, K. M. (2021). Variant-specific inflation factors for assessing population stratification at the phenotypic variance level. *Nature Communications*, *12*(1), 3506. https://doi.org/10.1038/s41467-021-23655-2

Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., Ibrahim, A., Ji, Y., John, S., Lewis, E., MacArthur, J. A. L., McMahon, A., Osumi-Sutherland, D., Panoutsopoulou, K., Pendlington, Z., … Harris, L. W. (2023). The NHGRI-EBI GWAS Catalog: Knowledgebase and deposition resource. *Nucleic Acids Research*, *51*(D1), D977–D985. https://doi.org/10.1093/nar/gkac1010

Sutter, N. B., Bustamante, C. D., Chase, K., Gray, M. M., Zhao, K., Zhu, L., Padhukasahasram, B., Karlins, E., Davis, S., Jones, P. G., Quignon, P., Johnson, G. S., Parker, H. G., Fretwell, N., Mosher, D. S., Lawler, D. F., Satyaraj, E., Nordborg, M., Lark, K. G., … Ostrander, E. A. (2007). A single *IGF1* allele is a major determinant of small size in dogs. *Science*, *316*(5821), 112–115. https://doi.org/10.1126/science.1137045

Swinford, N., Gallagher, B., Sheehama, J., Lin, M., Prall, S., Scelza, B., & Henn, B. M. (2023). Examination of runs of homozygosity in relation to height in an endogamous Namibian population. *American Journal of Biological Anthropology*, *180*(1), 207–215. https://doi.org/10.1002/ajpa.24660

Szpiech, Z. A., Blant, A., & Pemberton, T. J. (2017). *GARLIC*: Genomic autozygosity regions likelihood-based inference and classification. *Bioinformatics*, *33*(13), 2059–2062. https://doi.org/10.1093/bioinformatics/btx102

Tachmazidou, I., Süveges, D., Min, J. L., Ritchie, G. R. S., Steinberg, J., Walter, K., Iotchkova, V., Schwartzentruber, J., Huang, J., Memari, Y., McCarthy, S., Crawford, A. A., Bombieri, C., Cocca, M., Farmaki, A.-E., Gaunt, T. R., Jousilahti, P., Kooijman, M. N., Lehne, B., … Zeggini, E. (2017). Whole-genome sequencing coupled to imputation discovers genetic signals for anthropometric traits. *The American Journal of Human Genetics*, *100*(6), 865–884. https://doi.org/10.1016/j.ajhg.2017.04.014

Tchernov, E., & Valla, F. F. (1997). Two new dogs, and other Natufian dogs, from the Southern Levant. *Journal of Archaeological Science*, *24*(1), 65–95. https://doi.org/10.1006/jasc.1995.0096

Thompson, E. A. (2013). Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations. *Genetics*, *194*(2), 301–326. https://doi.org/10.1534/genetics.112.148825

Trujillo-Viera, J., El-Merahbi, R., Schmidt, V., Karwen, T., Loza-Valdes, A., Strohmeyer, A., Reuter, S., Noh, M., Wit, M., Hawro, I., Mocek, S., Fey, C., Mayer, A. E., Löffler, M. C., Wilhelmi, I., Metzger, M., Ishikawa, E., Yamasaki, S., Rau, M., … Sumara, G. (2021). Protein Kinase D2 drives chylomicron-mediated lipid transport in the intestine and promotes obesity. *EMBO Molecular Medicine*, *13*(5), e13548. https://doi.org/10.15252/emmm.202013548

Turner, S. D. (2014). *qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots*. https://doi.org/10.1101/005165

Urfer, S. R., Wang, M., Yang, M., Lund, E. M., & Lefebvre, S. L. (2019). Risk factors associated with lifespan in pet dogs evaluated in primary care veterinary hospitals. *Journal of the American Animal Hospital Association*, *55*(3), 130–137. https://doi.org/10.5326/JAAHA-MS-6763

Van Den Berg, S., Vandenplas, J., Van Eeuwijk, F. A., Lopes, M. S., & Veerkamp, R. F. (2019). Significance testing and genomic inflation factor using high-density genotypes or whole-genome sequence data. *Journal of Animal Breeding and Genetics*, *136*(6), 418–429. https://doi.org/10.1111/jbg.12419

Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall, T., Seppälä, E. H., Hansen, M. S. T., Lawley, C. T., Karlsson, E. K., The LUPA Consortium, Bannasch, D., Vilà, C., Lohi, H., Galibert, F., Fredholm, M., Häggström, J., Hedhammar, Å., … Webster, M. T. (2011). Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genetics*,

1045 *7*(10), e1002316. https://doi.org/10.1371/journal.pgen.1002316
1046 Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., Crandall,
1047 K. A., Lundeberg, J., & Wayne, R. K. (1997). Multiple and ancient origins of the domestic
1048 dog. *Science*, *276*(5319), 1687–1689. https://doi.org/10.1126/science.276.5319.1687
1049 vonHoldt, B. M., Pollinger, J. P., Lohmueller, K. E., Han, E., Parker, H. G., Quignon, P.,
1050 Degenhardt, J. D., Boyko, A. R., Earl, D. A., Auton, A., Reynolds, A., Bryc, K., Brisbin,
1051 A., Knowles, J. C., Mosher, D. S., Spady, T. C., Elkahloun, A., Geffen, E., Pilot, M., …
1052 Wayne, R. K. (2010). Genome-wide SNP and haplotype analyses reveal a rich history
1053 underlying dog domestication. *Nature*, *464*(7290), 898–902.
1054 https://doi.org/10.1038/nature08837
1055 Wang, C., Wallerman, O., Arendt, M.-L., Sundström, E., Karlsson, Å., Nordin, J., Mäkeläinen,
1056 S., Pielberg, G. R., Hanson, J., Ohlsson, Å., Saellström, S., Rönnberg, H., Ljungvall, I.,
1057 Häggström, J., Bergström, T. F., Hedhammar, Å., Meadows, J. R. S., & Lindblad-Toh, K.
1058 (2021). A novel canine reference genome resolves genomic architecture and uncovers
1059 transcript complexity. *Communications Biology*, *4*(1), 185.
1060 https://doi.org/10.1038/s42003-021-01698-x
1061 Wang, G.-D., Zhai, W., Yang, H.-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu,
1062 C.-L., Poyarkov, A. D., Irwin, D. M., Hytönen, M. K., Lohi, H., Wu, C.-I., Savolainen, P., &
1063 Zhang, Y.-P. (2016). Out of southern East Asia: The natural history of domestic dogs
1064 across the world. *Cell Research*, *26*(1), 21–33. https://doi.org/10.1038/cr.2015.147
1065 Warrington, N. M., Beaumont, R. N., Horikoshi, M., Day, F. R., Helgeland, Ø., Laurin, C.,
1066 Bacelis, J., Peng, S., Hao, K., Feenstra, B., Wood, A. R., Mahajan, A., Tyrrell, J.,
1067 Robertson, N. R., Rayner, N. W., Qiao, Z., Moen, G.-H., Vaudel, M., Marsit, C. J., …
1068 Freathy, R. M. (2019). Maternal and fetal genetic effects on birth weight and their
1069 relevance to cardio-metabolic risk factors. *Nature Genetics*, *51*(5), 804–814.
1070 https://doi.org/10.1038/s41588-019-0403-1

Wayne, R. K. (1993). Molecular evolution of the dog family. *Trends in Genetics*, *9*(6), 218–224. https://doi.org/10.1016/0168-9525(93)90122-X

Wu, W., Li, D., Feng, X., Zhao, F., Li, C., Zheng, S., & Lyu, J. (2021). A pan-cancer study of selenoprotein genes as promising targets for cancer therapy. *BMC Medical Genomics*, *14*(1), 78. https://doi.org/10.1186/s12920-021-00930-1

Yang, Q., Chen, H., Ye, J., Liu, C., Wei, R., Chen, C., & Huang, L. (2019). Genetic diversity and signatures of selection in 15 Chinese indigenous dog breeds revealed by genome-wide SNPs. *Frontiers in Genetics*, *10*, 1174. https://doi.org/10.3389/fgene.2019.01174

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian, X., Wang, G., Luo, Q., Zhang, Q., Liu, Q., & Xiong, L. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*, *5*(1), 5087. https://doi.org/10.1038/ncomms6087

Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliasen, A. U., Jiang, Y., Raghavan, S., Miao, J., Arias, J. D., Graham, S. E., Mukamel, R. E., Spracklen, C. N., Yin, X., Chen, S.-H., Ferreira, T., Highland, H. H., … Hirschhorn, J. N. (2022). A saturated map of common genetic variants associated with human height. *Nature*, *610*(7933), 704–712. https://doi.org/10.1038/s41586-022-05275-y

Yordy, J., Kraus, C., Hayward, J. J., White, M. E., Shannon, L. M., Creevy, K. E., Promislow, D. E. L., & Boyko, A. R. (2020). Body size, inbreeding, and lifespan in domestic dogs. *Conservation Genetics*, *21*(1), 137–148. https://doi.org/10.1007/s10592-019-01240-x

Yuan, J., Wang, X., Zhao, Y., Khan, N. U., Zhao, Z., Zhang, Y., Wen, X., Tang, F., Wang, F., & Li, Z. (2020). Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. *Scientific Reports*, *10*(1), 9958. https://doi.org/10.1038/s41598-020-66604-7

Zhao, W., Yang, H., Chai, J., & Xing, L. (2021). RUNX2 as a promising therapeutic target for malignant tumors. *Cancer Management and Research*, *Volume 13*, 2539–2548.

https://doi.org/10.2147/CMAR.S302173